

Open Research Online

The Open University's repository of research publications and other research outputs

Information-seeking on the Web with Trusted Social Networks - from Theory to Systems

Thesis

How to cite:

Heath, Tom (2008). Information-seeking on the Web with Trusted Social Networks - from Theory to Systems. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 2008 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000ea56>

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

oro.open.ac.uk

Information-seeking on the Web with Trusted Social Networks – from Theory to Systems

Tom Heath, BSc

Knowledge Media Institute

The Open University

Thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Computer Science

Submitted 14th January 2008

Date of award 28 Oct. 2008

Abstract

This research investigates how synergies between the Web and social networks can enhance the process of obtaining relevant and trustworthy information. A review of literature on personalised search, social search, recommender systems, social networks and trust propagation reveals limitations of existing technology in areas such as relevance, collaboration, task-adaptivity and trust.

In response to these limitations I present a Web-based approach to information-seeking using social networks. This approach takes a source-centric perspective on the information-seeking process, aiming to identify trustworthy sources of relevant information from within the user's social network.

An empirical study of source-selection decisions in information- and recommendation-seeking identified five factors that influence the choice of source, and its perceived trustworthiness. The priority given to each of these factors was found to vary according to the criticality and subjectivity of the task.

A series of algorithms have been developed that operationalise three of these factors (*expertise, experience, affinity*) and generate from various data sources a number of trust metrics for use in social network-based information seeking. The most significant of these data sources is *Revyu.com*, a reviewing and rating Web site implemented as part of this research, that takes input from regular users and makes it available on the Semantic Web for easy re-use by the implemented algorithms.

Output of the algorithms is used in *Hoonoh.com*, a Semantic Web-based system that has been developed to support users in identifying relevant and trustworthy information

sources within their social networks. Evaluation of this system's ability to predict source selections showed more promising results for the *experience* factor than for *expertise* or *affinity*. This may be attributed to the greater demands these two factors place in terms of input data. Limitations of the work and opportunities for future research are discussed.

Acknowledgements

Many people have helped make this research possible; adequately acknowledging them has been one of the most challenging writing tasks of this dissertation.

Starting at the beginning, an inspired idea from Mo Heath circa 1995 gave me my first proper experience of the Web, and I caught the bug. At various times since then Susan Martin, Andy Green, Peter Goodhew and the Fenton Group all helped create environments (real or virtual) in which I could develop the technical skills and understanding that underpins some of the more practical parts of this research.

Martin Poulter first introduced me to the FOAF vocabulary in 2001, and by extension the Semantic Web. At this point I caught another bug. This one took slightly longer to properly understand, and I owe many thanks to all the people who have helped relieve me of my ignorance along the way.

In 2004 Steve Richardson gave me a recommendation that demonstrated the power of affinity before I even knew the concept existed, and brought home to me the urgent need for a social element to information-seeking on the Web. Thanks to Steve and the staff of The Atlanta in Bangkok for the inspiration.

Simon Chu's supervision of my undergraduate research project and, strangely enough, his accounts of collecting sweaty t-shirts in the name of science helped convince me that academic research was a worthwhile and interesting pursuit. Matt Murphy provided valuable guidance when I decided to take the plunge.

A chance encounter with Marc Eisenstadt over Jabber ultimately led me to KMi and The Open University. The KMi 'vibe' (there's no word that better sums it up), for which Marc

deserves so much of the credit, makes for an institute where doing cutting-edge research is not only possible, but a real pleasure. Marc has provided ongoing inspiration, guidance and constructively critical input for which I'm truly grateful. From him I take many things, not least of which a constant reminder to ask "so what?!"

In my time at the lab I've been fortunate in having what I believe to be a truly world-class team of people supervising my work: Enrico Motta, Marian Petre and Martin Dzbor.

Whilst different in many ways, Enrico and Marian share a hawk-like ability to spot a flawed narrative a mile off, which is both humbling and awesome to watch. They've made a great team.

Enrico has been a first-class source of wisdom, insight and encouragement, and I feel incredibly fortunate to have been his student. In addition to providing expert technical guidance, Enrico has always encouraged me to take the opportunities that have made this PhD so rewarding.

I still don't think I've convinced Marian of the wonders the Semantic Web has to offer, but trying has certainly forced me to clarify my own ideas. Being supervised by one of the authors of what must be the funniest and most insightful book about doing PhD research is a true privilege.

Martin Dzbor provided invaluable guidance during my application process, followed by substantial opportunities for exploration of ideas during the early part of the PhD.

I owe thanks to many other people within KMi. Together with Marian, Trevor Collins runs the weekly Postgraduate Forum training sessions within The Open University's Centre for Research in Computing. These events don't just provide caffeine, chocolate

brownies and rice crispy cakes; they are the hub of a wonderful community of PhD students across the CRC and have played a major part in shaping me as a researcher. I'm immensely grateful to Trevor and Marian for their commitment, wisdom and stamina.

John Domingue took a gamble and gave me the opportunity to serve on the Organising Committee of ESWC2006 as Semantic Web Technologies Coordinator. This was one of the most stressful yet rewarding things I've ever done, and opened many doors; thanks John.

The administrative and technical teams within KMi (Ortenz, Jane, Catherine, Aneta, Paul, Damian, Lewis and Robbie) are first class and deserve huge thanks. The Open University as a whole has proved to be a fantastic place to do a PhD. I've met great people, and feel honoured to be a part of this unique institution. Further afield, the EPSRC-funded Advanced Knowledge Technologies (AKT) project helped fund this PhD, and provided access to a ready-made research community across the UK.

Steve Cayzer and all members of the Semantic and Adaptive Systems Department at HP Labs Bristol allowed me to camp out in their rich and stimulating environment whilst I cut my teeth on Semantic Web technologies and did early work on Revyu.com. Many people have contributed reviews to Revyu; Paddy, drewp, hockeyshooter and martinp deserve special mentions. I'm also grateful to all the people who gave up their time to participate in the two studies reported in this dissertation. Peter Coetzee wrote the crawler that collected the FOAF data used in Hoonoh.com. Using a data set of this size is possible thanks to the Talis Platform.

Dinar Kale told me on day one that doing a PhD is a marathon, not a sprint. For this pearl of wisdom I thank him sincerely. If there is one piece of advice I would pass on to future

doctoral students this is it. Debra Haley has been a great friend and a writing buddy extraordinaire. Without her the process would have been far more painful and much less fun.

Sue and Rog have been wonderfully willing listeners throughout the entire PhD, and have been kind enough to share with me their wonderful, distraction-free home, where large portions of this dissertation were written.

The Heath family has been a constant source of inspiration since long before this PhD. Most of all I thank them for being who they are, and for giving me the confidence and self-belief that has seen me through.

Lastly I'm endlessly grateful to Livs, the greatest co-pilot I could wish for, who in addition to providing relentless encouragement and cheerleading has helped ensure that not all my meals over the last few months have been served in a bowl.

Publications

The research reported in this dissertation has contributed to the following publications:

- Tom Heath and Enrico Motta (2008) Ease of Interaction plus Ease of Integration: Combining Web2.0 and the Semantic Web in a Reviewing Site. *Journal of Web Semantics*, 6 (1) (Special Issue on Web2.0 and the Semantic Web).
- Tom Heath and Enrico Motta (2007) Revyu.com: A Reviewing and Rating Site for the Web of Data. In *Proceedings of the 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference (ISWC+ASWC2007)*, Busan, Korea. LNCS 4825.
- Tom Heath, Enrico Motta and Marian Petre (2007) Computing Word-of-Mouth Trust Relationships in Social Networks from Semantic Web and Web2.0 Data Sources. In *Proceedings of the Workshop on Bridging the Gap between Semantic Web and Web 2.0, 4th European Semantic Web Conference (ESWC2007)*, Innsbruck, Austria.
- Tom Heath, Enrico Motta (2006) Reviews and Ratings on the Semantic Web. In *Poster Track, 5th International Semantic Web Conference (ISWC2006)*, Athens, Georgia, USA.
- Tom Heath and Enrico Motta (2006) Personalizing Relevance on the Semantic Web through Trusted Recommendations from a Social Network. In *Proceedings of the ESWC2006 Workshop on Semantic Web Personalization*, Budva, Montenegro.

- Tom Heath, Enrico Motta, and Marian Petre (2006) Person to Person Trust Factors in Word of Mouth Recommendation. In Proceedings of the CHI2006 Workshop on Reinventing Trust, Collaboration, and Compliance in Social Systems (Reinvent06), Montréal, Canada.
- Tom Heath, Enrico Motta, and Martin Dzbor (2005) Context as Foundation for a Semantic Desktop. In Proceedings of the 1st Workshop on the Semantic Desktop, International Semantic Web Conference (ISWC2005), Galway, Ireland.
- Tom Heath, Enrico Motta, and Martin Dzbor (2005) Uses of Contextual Information to Support Online Tasks. In Proceedings of the 1st AKT Doctoral Symposium, Milton Keynes, UK.
- Tom Heath, Martin Dzbor, and Enrico Motta (2005) Supporting User Tasks and Context: Challenges for Semantic Web Research. In Proceedings of the Workshop on End-user Aspects of the Semantic Web (UserSWeb), European Semantic Web Conference (ESWC2005), Heraklion, Crete.
- Tom Heath (2005) The Modelling, Capture, and Use of Social Context in Online Tasks. KMi Technical Report: kmi-05-10, September 2005.

Contents

Abstract..... 2

Acknowledgements 4

Publications 8

Contents 10

List of Figures..... 14

List of Tables 16

List of Quotes 17

List of Code Fragments 18

1. Introduction..... 19

1.1. Motivation..... 19

1.2. Approach..... 20

1.3. Research Questions..... 21

1.4. Definition of Terms..... 22

1.5. Structure of the Dissertation 23

2. Literature Review 25

2.1. Information-seeking on the Web 26

2.2. Relevance: Topical vs. Personal 27

2.3. Personalised Search 29

2.4. Information Filtering with 'Classic' Recommender Systems..... 31

2.4.1. Content-Based Recommendation 33

2.4.2. Collaborative Filtering..... 33

2.5. Social Navigation..... 39

2.6. Approaches Based on Social Networks 42

2.6.1. Social Search (Network) and Expert Finding 42

2.6.2. Social Networks and Trust..... 43

2.7. Conclusions and Gap Analysis 49

2.7.1. The Nature of Relevance 50

2.7.2. Economies of Scale through Collaboration 50

2.7.3. Personalisation and Recommendation across Varied Domains..... 51

2.7.4. Personalisation and Recommendation in Open Worlds 52

2.7.5. Sparsity and Cold-Start Problems..... 53

2.7.6. Richness of Trust Models 54

2.7.7. Summary..... 56

3. Approach: Personalised Relevance in Information-seeking through a Trusted Social Network 58

3.1. Characteristics of this Approach..... 59

- 3.1.1. Source-centricity 59
 - 3.1.2. Task-adaptivity..... 60
- 3.2. Benefits of this Approach..... 62
 - 3.2.1. Increased Personal Relevance 62
 - 3.2.2. Utility across a Range of Tasks..... 64
 - 3.2.3. Spam-resistant Information-seeking 64
 - 3.2.4. Openness to Additional Information..... 65
 - 3.2.5. Greater Reuse of Existing Data..... 67
- 3.3. Summary 68
- 4. Source Selection in Word-of-mouth Information-seeking 70**
 - 4.1. Background and Related Work 70
 - 4.1.1. The Role of Weak Ties 71
 - 4.1.2. The Role of Strong Ties 72
 - 4.1.3. Influences on Choice of Tie-Strength 74
 - 4.1.4. The Role of Source, Task and Individual Characteristics..... 76
 - 4.1.5. Summary 83
 - 4.2. Study of Source Selection in Word-of-mouth Information-seeking 86
 - 4.3. Design 86
 - 4.3.1. Pilot 87
 - 4.4. Method 87
 - 4.4.1. Participants..... 87
 - 4.4.2. Procedure 87
 - 4.5. Analysis..... 91
 - 4.6. Results and Discussion..... 92
 - 4.6.1. Trust Factors: Definitions 92
 - 4.6.2. Trust Factors: Illustrative Quotes..... 93
 - 4.6.3. Trust Factors: Occurrence Frequencies..... 96
 - 4.6.4. Trust Topicality 97
 - 4.6.5. Relation of Trust Factors to Previous Work 97
 - 4.6.6. General Trends in Application of Trust Factors 99
 - 4.6.7. Effects of Criticality, Subjectivity 99
 - 4.6.8. Effects of Task Modality..... 100
 - 4.6.9. Domain of Task and Nature of Relationship 101
 - 4.7. Conclusions 101
- 5. Technical Approach and Architecture..... 103**
 - 5.1. Architectural Overview 103
 - 5.2. Computing Trust Relationships..... 105
 - 5.2.1. Choice of Trust Factors..... 106
 - 5.2.2. Data Acquisition from Distributed Sources 107
 - 5.3. The Semantic Web 110

- 5.3.1. Resource Description Framework (RDF) 111
 - 5.3.2. Using RDF for Data Integration 112
 - 5.3.3. Linked Data..... 114
 - 5.3.4. Querying RDF Data 115
 - 5.3.5. Summary 116
- 5.4. Sources of Social Network Data 117
- 5.5. Implemented Systems and Detailed Architecture..... 119
- 6. Revyu: a Semantic Web Reviewing and Rating Site 121**
 - 6.1. Introduction..... 121
 - 6.2. Novel Features of Revyu 122
 - 6.3. Related Work 124
 - 6.4. User Walkthrough..... 126
 - 6.4.1. Generating Semantic Web Content by Completing Web Forms 127
 - 6.5. Use of Common Ontologies..... 130
 - 6.6. The Role of Tagging in Revyu..... 132
 - 6.6.1. Tagging versus Classification..... 132
 - 6.6.2. Inferring the Type of Reviewed Items 134
 - 6.6.3. Identifying Related Tags..... 135
 - 6.7. The Role of Links in Identifying Reviewed Items..... 136
 - 6.8. Technical Implementation 138
 - 6.9. Linked Data Compliance 139
 - 6.10. Links to other Data Sets..... 142
 - 6.11. Consuming Linked Data 143
 - 6.11.1. Supplementing Reviewer Information with FOAF Data 144
 - 6.11.2. Supplementing Film and Book Reviews with External Data 145
 - 6.11.3. Supplementing Reviewed Items by Pre-population..... 149
 - 6.12. Reusability of Revyu Data 152
 - 6.13. Conclusions..... 153
- 7. Hoonoh: Source-centric Information-seeking with Trusted Social Networks .. 155**
 - 7.1. Introduction..... 155
 - 7.2. Computing Knowledge and Trust Relationships: The Hoonoh Algorithms..... 156
 - 7.2.1. Input Data to Hoonoh Algorithms 157
 - 7.2.2. Tags as Topics 157
 - 7.2.3. Proxy Metrics..... 158
 - 7.2.4. Computing Trust Metrics from Revyu Data 161
 - 7.2.5. Supplementing Experience Metrics through Additional Data Sources 170
 - 7.3. Technical Implementation of the Hoonoh Algorithms 173
 - 7.4. Representing Computed Trust Relationships..... 175
 - 7.5. Hoonoh..... 180
 - 7.5.1. Hoonoh from the User Perspective 180

- 7.5.2. Ranking by Trust Factors 183
- 7.5.3. Filtering by Social Network 185
- 7.5.4. Browsing Functionality 187
- 7.5.5. Queries for Multiple Topics 188
- 7.5.6. System Implementation 190
- 7.5.7. Summary 191
- 8. Evaluation 192**
 - 8.1. Introduction 192
 - 8.2. Design 193
 - 8.3. Method 194
 - 8.3.1. Participants 194
 - 8.3.2. Procedure 195
 - 8.4. Analysis 199
 - 8.5. Results 202
 - 8.6. Discussion 205
- 9. Conclusions and Outlook 210**
 - 9.1. Summary of the Research 210
 - 9.2. Contributions of the Research 211
 - 9.2.1. Contribution 1 211
 - 9.2.2. Contribution 2 212
 - 9.2.3. Contribution 3 213
 - 9.2.4. Contribution 4 213
 - 9.2.5. Summary 214
 - 9.3. Limitations and Future Work 214
 - 9.3.1. Availability of Data 215
 - 9.3.2. Scale 222
 - 9.3.3. Reliance on Proxy Metrics 224
 - 9.3.4. Exploiting Semantics for Trust Generation 225
 - 9.3.5. Combining Trust Metrics 228
 - 9.3.6. Evaluation Methods 229
 - 9.4. Outlook 229
- References 234**
- Appendices 243**
- Appendix A: Interview Script 244**
- Appendix B: Master list of themes and superordinate themes identified in the interview study 247**

List of Figures

Figure 1. Conceptual structure of the literature review, showing limitations of and links between approaches..... 25

Figure 2. The information-seeking process from a source identification and source selection perspective..... 84

Figure 3. Response frequencies for each factor, summed across 4 scenarios in each of 12 interviews 97

Figure 4. Overview of the technical architecture..... 104

Figure 5. Architecture of systems developed to support, or used within, this research.. 120

Figure 6. The Revyu Home Page..... 122

Figure 7. The HTML view of a review on Revyu 127

Figure 8. The upper half of the Revyu Reviewing Form..... 128

Figure 9. Example of RDF output of a Revyu review 131

Figure 10. The lower half of the Revyu reviewing form 137

Figure 11. Links from Revyu to external data sets 143

Figure 12. The author's Revyu profile page, showing review data from the site (left) alongside information from his external FOAF file (right)..... 145

Figure 13. A film review on Revyu (left) shown alongside film data from DBpedia (right)..... 147

Figure 14. A book review on Revyu (left) shown alongside film data from Amazon/RDF Book Mashup (right) 148

Figure 15. Geodata from the Open Guide to Milton Keynes used to display a Map of a reviewed item's location 151

Figure 16. Revyu review data reused in the Semantic Web gateway *Watson*, via the Revyu SPARQL endpoint 153

Figure 17. Usage (Experience) algorithm in pseudo-code..... 162

Figure 18. Credibility (Expertise) algorithm in pseudo-code 165

Figure 19. Affinity algorithm in pseudo-code 168

Figure 20. Schematic diagram showing the relationships between classes and properties in the Hoonoh ontology 177

Figure 21. The Hoonoh Home Page 180

Figure 22. Example output: the author's social network ranked by expertise in the topic 'film' 181

Figure 23. Example output: the author's social network ranked by a combination of experience in the topic 'film' and affinity 182

Figure 24. Example output: detail pages such as this show items the information source has reviewed that relate to the topic of the query. 183

Figure 25. Weighting options in the Hoonoh results interface 1

Figure 26. Example output: weighted results for anonymous users 1

Figure 27. The 'topics' area of the site allows users to browse for potential information sources by topic 1

Figure 28. The 'people' area of the site allows users to browse the topics in which specific individuals have experience or expertise. Logged in users will be shown people who the system knows to be in their network..... 1

Figure 29. Hoonoh system architecture 1

List of Tables

Table 1. Source and relational factors identified in existing literature as affecting perceived information quality..... 85

Table 2. Recommendation-seeking scenarios used in interviews with participants 89

Table 3. Distribution of usage (experience) scores in a sample Hoonoh data set 164

Table 4. Distribution of credibility (expertise) scores in a sample Hoonoh data set 166

Table 5. Distribution of affinity scores in a sample Hoonoh data set..... 170

Table 6. Frequency of reviews created by each participant in the evaluation 197

Table 7. Combinations of trust factors and weightings used to generate ranked results 200

Table 8. Spearman rank coefficients of correlation between participant responses and Hoonoh output with each ranking combination. Entries in bold are significant at the 0.05 alpha level..... 203

List of Quotes

Quote 1. Participant ID 16, Plumber scenario.....

Quote 2. Participant ID 10, Back Pain scenario.....

Quote 3. Participant ID 05, Plumber scenario.....

Quote 4. Participant ID 14, Plumber scenario.....

Quote 5. Participant ID 08, Holiday Activities scenario.....

Quote 6. Participant ID 16, Holiday Activities scenario.....

Quote 7. Participant ID 17, Holiday Activities scenario.....

Quote 8. Participant ID 12, Plumber scenario.....

Quote 9. Participant ID 07, Back Pain scenario.....

Quote 10. Participant ID 16, Hotel scenario

Quote 11. Participant ID 12, Holiday Activities scenario..... 1

List of Code Fragments

Code Fragment 1. Example of the structure of a CustomerFull Response Group to a CustomerContentLookup Operation on the Amazon Associates Web Service API. 109

Code Fragment 2. The author's Del.icio.us social network as a JSON array, retrieved from the Del.icio.us API at <http://del.icio.us/feeds/json/network/tomheath> 110

Code Fragment 3. Example RDF document demonstrating the use of the `foaf : name` and `foaf : homepage` properties 112

Code Fragment 4. Example RDF document showing the use of `owl : sameAs` to state that a number of different URIs represent the same resource 115

Code Fragment 5. An example Expertise relationship described using the Hoonoh ontology 178

Code Fragment 6. An example Affinity relationship described using the Hoonoh ontology 179

1. Introduction

1.1. *Motivation*

Word-of-mouth is a powerful mechanism for obtaining information. Using one's known social network as an information source enables detailed assessments to be made about the potential relevance and trustworthiness of the information received (Kautz, Selman and Shah, 1997b, 1997a). However, the current generation of Web tools is not well adapted to these mechanisms, and consequently tasks that rely on these tools do not gain the benefits of well-established social processes. This leaves a disparity between the strategies used for *information-seeking* (Marchionini, 1995) in online and offline environments.

The World Wide Web is a global information space on an unprecedented scale, made up of interlinked documents so great in number that they push the boundaries of human comprehension. Searching for information is one of the most popular activities on the Web (Haythornthwaite, 2001), and without text search engines such as Google that index billions of Web documents our access to this immense information resource would be severely limited. The very scale of the Web means that thousands of results can be returned by a search engine in response to one query. These results are typically ranked for relevance to the search terms according to the keywords they contain and a coarse notion of 'popularity' indicated by the number of incoming links a document receives (Page, Brin, Motwani et al., 1999).

Despite the power and complexity of relevance ranking algorithms, there is great potential for *information overload* (Denning, 2006). This is exacerbated by the 'one-size-

fits-all' nature of ranking algorithms, where two users would receive the same results for a query irrespective of whether their underlying information needs were the same (Teevan, Dumais and Horvitz, 2005). Search engine results pages typically provide few additional cues on which the user can judge the relevance of results to his or her information need. This can be particularly problematic in cases that require careful judgement, selection of information from trusted sources, or where relevance depends heavily on user-specific needs and preferences.

1.2. Approach

This dissertation explores an approach to information seeking on the Web, based on the principles of word-of-mouth recommendation in social networks. Word-of-mouth through traditional channels (e.g. voice, face-to-face) provides access to new information that would not otherwise be available to the information seeker, and helps to filter out less relevant items from a broader pool of options (Granovetter, 1973, Kautz, Selman et al., 1997b, 1997a). This research explores these mechanisms in more detail and applies them in the virtual world.

The adopted approach is oriented around first identifying the most appropriate and trusted sources of recommendations and then using the knowledge held by these individuals to assist in the information seeking task. By combining technical systems that harness the knowledge and experience of users' social networks with their own knowledge of members of those networks, the goal is to reduce information overload and provide access to information that is more personally relevant and trustworthy.

This dissertation addresses the following research questions.

1.3. *Research Questions*

The research reported in this dissertation addresses the following principal question:

'To what extent can information- and recommendation-seeking within social networks be supported on the Web?'

This question can be broken down into a number of specific research questions:

1. How do people choose information and recommendation sources from among members of their social network?
2. Which factors influence judgements about the relevance and trustworthiness of these information and recommendation sources?
3. How do the characteristics of the task being performed affect these judgements?
4. To what extent can general principles derived from answers to the previous questions be operationalised as computational algorithms that replicate the process of seeking information and recommendations through social networks?
5. How feasible is the implementation of user-oriented systems that exploit such algorithms?
6. If such systems can be implemented, how do they perform relative to human performance of equivalent tasks?

1.4. Definition of Terms

Specialist terms, or those whose meaning may be open to interpretation, will mostly be defined in the body of the dissertation as required. However, for the sake of clarity a number of terms will be defined at this stage.

Information seeking is "a process in which humans purposefully engage in order to change their state of knowledge" (Marchionini, 1995) (pp. 5). This dissertation does not treat *information-seeking* and *recommendation-seeking* as distinct processes but as variations on the same theme. Recommendation-seeking is seen simply as a form of information-seeking in which the information seeker tries to obtain opinions or value judgements from trusted (or otherwise favoured) sources, as a means to distinguish between potentially relevant items and thereby reduce the search space.

In the context of this research, an individual's *social network* is defined in the first instance as the people they know personally and with whom they identify in some way, possibly through shared characteristics, socio-cultural identity, or other group membership. This may encompass family members, friends, colleagues, neighbours or other acquaintances. In the second instance the definition of the network may be extended to take in so-called 'friends-of-friends' (those people in the networks of members of one's own network), or even 'friends-of-friends-of-friends'.

This research views both social networks and *relevance* as being primarily a construction of the individual. On this basis, no assumptions are made from the outset about how the nature or origin of social relations may influence the information-seeking process. In particular, no assumptions are made about how particular classes of network members

might contribute to the information-seeking process, as these issues will be examined by research presented later in this dissertation.

1.5. *Structure of the Dissertation*

The following chapter (2) presents a thorough review of related work, touching on fields such as *relevance*, *social navigation* and *recommender systems*. Being concerned with information-seeking through the medium of the Web, the research acknowledges much of the prior work on Web search engines and information retrieval but does not examine this in detail.

From this review gaps in existing research are identified. These inform the approach pursued in this research, which is outlined in Chapter 3.

Chapter 4 reviews additional literature related to information-seeking via word-of-mouth, before reporting on an empirical study of how people choose sources of information and recommendations from within their social network and make judgements about the trustworthiness of these sources. This study yields novel results regarding the factors that influence source selection when seeking information from one's social network and patterns in how these are applied across different tasks. These findings also highlight how the characteristics of the information-seeking task can influence source selection.

Chapter 5 outlines the technical approach and distributed architecture adopted in this research, through which the theoretical principles are instantiated and further investigated. Chapter 6 presents the award-winning¹ reviewing and rating Web site

¹ Revyu.com was awarded first prize in the 2007 Semantic Web Challenge

Revyu.com, the first of two Semantic Web-based applications that have been implemented as part of this research and contribute to the broader architecture.

In Chapter 7 a methodology and algorithms are presented for deriving metrics that describe trust relationships in social network-based information-seeking. These algorithms have been developed based on the findings from the empirical study presented in Chapter 4, and operate on data from Revyu and a range of other Web2.0 and Semantic Web data sources. Description of the algorithms is followed by presentation of Hoonoh.com, a live, publicly accessible Web site based directly on the principles and findings of this research. Hoonoh is a demonstration of how technical systems can assist people in seeking information supported by their trusted social networks.

Chapter 8 reports on a study evaluating the effectiveness of the Hoonoh algorithms in predicting individuals' choice of information sources. Limitations of the research are discussed in Chapter 9, alongside identification of future directions for this research.

2. Literature Review

This chapter reviews literature in the fields of information-seeking, relevance, personalised search, recommender systems, social navigation, social search and trust. The structure of the review is outlined in Figure 1 below, which also illustrates interrelations between these fields and how the limitations of work in one field motivate related work in another.

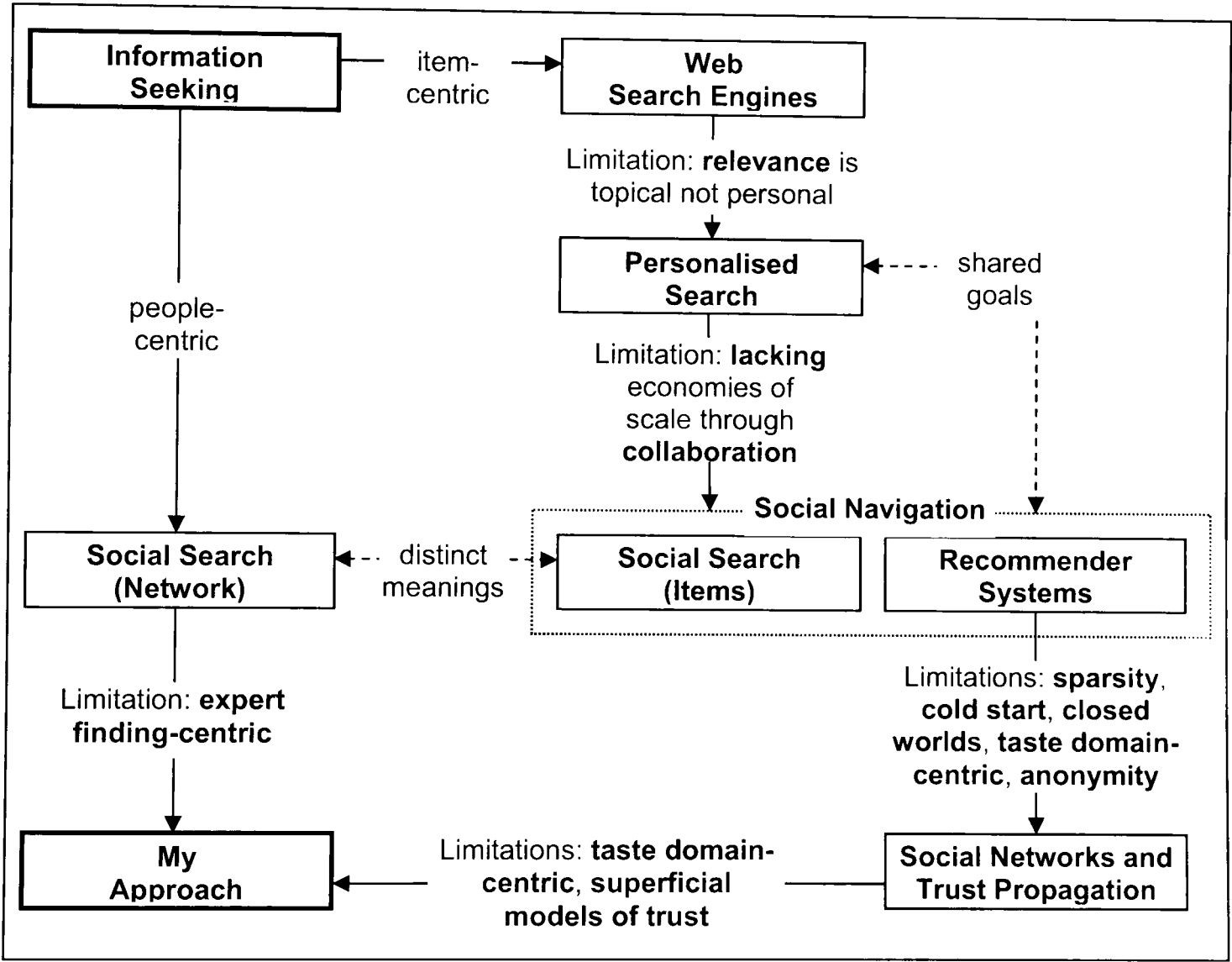


Figure 1. Conceptual structure of the literature review, showing limitations of and links between approaches

The limitations shown in Figure 1 are discussed throughout this chapter; pointers to the relevant sections are provided here for the reader's convenience: *relevance* (2.2, 2.7.1), *lacking collaboration* (2.3, 2.7.2), *sparsity* and *cold-start* (2.6.2, 2.7.5), *closed worlds*

(2.7.4), *taste domain-centricity* (2.7.3), *anonymity* and *superficial models of trust* (2.7.6), and *expert-finding-centricity* (2.6.1).

2.1. Information-seeking on the Web

Where an individual encounters a problem or task for which their current knowledge is inadequate, they may engage in information-seeking in order to change their knowledge state (Belkin, 2000). Seeking information is one of the most common activities people perform on the Web (Haythornthwaite, 2001). Search engines such as Google attempt to support this process using complex algorithms that take account of the content of documents and their patterns of linkage to other documents (Page, Brin et al., 1999), in an attempt to identify documents that are most relevant to a user's search query.

However, despite (and often because of) the vast extent of online resources, locating the required piece of information can still present challenges to the user. These challenges may take a number of forms, for example:

- the user may not be able to identify suitable keywords that lead him to documents containing the information he requires, due to issues with synonymy (where two different terms are used to refer to the same concept) or polysemy (where one word has many meanings) (Narayanan, Koppaka, Edala et al., 2004), or because he is unsure of exactly what he is looking for (Belkin, 2000).
- the required information may not yet be available online because it is sensitive in nature (Kautz, Selman et al., 1997a) or stored in a legacy format or system.
- the user's search query may yield so many results that identifying those most relevant to the original information need is not always possible (Denning, 2006).

This last problem is an example of so-called *information overload* (Denning, 2006), where the vast numbers of documents on the Web that may be somewhat relevant to a query overwhelm the smaller subset of those of greatest relevance. In many cases, just one of the documents linked to from a search engine results page may be sufficient to meet the user's information need, but this document may not be easily identifiable among the many hundreds or thousands listed in the results. Individual human beings are not well equipped to quickly process and differentiate such large amounts of information, whereas machines are more adapted to this task.

These factors suggest that current Web search applications are inadequate in a number of scenarios, including when the user is unsure of exactly what they're looking for, or when a query may yield too many results. Furthermore, despite the vastness of the Web not all information is available through this medium, meaning that despite maintaining vast indices, search engines represent something of a closed world compared to the universe of human knowledge. Resolving these issues requires systems or processes that can identify potential sources of additional information that are not currently available on the Web, and more sophisticated means of filtering information based on its relevance to the individual's information needs.

2.2. Relevance: Topical vs. Personal

The relevance of results provided by any information-seeking system can be seen as a key factor in the system's effectiveness. Literature on information retrieval has traditionally viewed relevance as a measure of the suitability of a result to the information need of the user, *as that need is expressed in a query issued to the system.*

This relationship between document and query has been referred to as *topical relevance* (Eisenberg and Schamber, 1988).

Precision and *recall*, widely used measures of the effectiveness of information retrieval systems, are predicated on this notion of topical relevance and assume a closed corpus of documents over which a system may operate. This assumption is not sound if one considers the Web, and particularly human knowledge as a whole, as an open rather than closed world. Furthermore, by assuming some abstract notion of relevance these measures embody a positivist attitude to the information-seeking process (i.e. that there is a right answer or objective truth) at the expense of a more constructivist view. This may only be appropriate in domains where objectively correct solutions are more commonplace.

Authors such as Park (1994) (see also Kuhltau (1991)) have argued in favour of a notion of *personal relevance*, where the suitability of search results is considered relative to the abstract information need of the user, irrespective of how effectively this has been expressed in the search query. As Belkin (2000) argues, people can face significant problems "choosing the correct words to represent their information problems" (pp. 58). Therefore, whilst measuring personal relevance scientifically may prove challenging, simply measuring topical relevance without taking into account the user's task tells us relatively little about how well information systems are meeting user needs.

Building systems that enable personal relevance requires that additional knowledge about the user and their information needs, intent, task and context is taken into account. However, such information may be difficult to express via keyword search (Teevan,

Dumais et al., 2005), suggesting that additional techniques for representing such information in the system may be required.

2.3. *Personalised Search*

A number of researchers have attempted to offer personalised search, using a range of approaches for capturing broader information about the user from which to infer their information needs.

At the level of general Web search, Jeh and Widom (2003) present a modified version of the *PageRank* algorithm (Page, Brin et al., 1999). This approach takes as input a user's list of Web bookmarks, each of which is taken as an implicit endorsement of the relative importance of that document to the user. Based on this input, personalised *PageRank* scores can be calculated, which enables a personalised rather than a global view of the importance of Web documents, and can serve as a basis for ranking search results.

Specifically in the context of a job-seeking site, Bradley, Rafter and Smyth (2000) report on a system that filters search results by comparing these to a user profile based on jobs she has previously viewed and rated. This approach requires more extensive and ongoing input from the user compared to that of Jeh and Widom (2003), as the user must actively view and rate job advertisements in order to receive personalised results. The system is also domain-specific; however it could be extended to allow the capture of viewing and rating data for any corpus of items. It remains to be seen whether explicit ratings such as these, compared to the implicit endorsement of bookmarking a Web site, provide more accurate data on which to base user profiles for personalised search.

Teevan, Dumais et al. (2005) report on an investigation into the potential value of personalised search results compared to those provided by current search engines whereby all users receive the same results. They found that search results currently reflect a broad range of different search intents, meaning that relevance to the intents of the group as a whole is generally high. However, the relevance of generic results to individual search intentions was considerably lower. Interestingly it was found that agreement about the relevance of results between individuals choosing the same search query was lower than found in previous studies. This finding is attributed to the study's emphasis on participants rating the relevance of results to their personal information needs rather than an abstract notion of the results' relevance to a topic.

Furthermore, it was found that inter-rater agreement on the relevance of results was relatively low (62%) even for those participants who used the same query and whose expressed intentions were the same. It was concluded that participants struggled to unambiguously describe their search intention, therefore the same description actually covered more than one intention and the relevance rating of results varied as a consequence. Based on these findings, Teevan et al. conclude that there may be value in personalising search results, and propose a technical approach based on re-ordering results retrieved through conventional search engines.

Many search personalisation approaches are limited by only exploiting information provided specifically by that user. For example, Bradley, Rafter et al.'s (2000) system only bases personalisation on viewing and rating data from the user themselves. Similarly Jeh and Widom's (2003) approach does not specifically address the use of other people's bookmark collections to aid one's own personalisation.

Using data about just one user does not allow for economies of scale through collaboration, whereby the knowledge and experience of other people could be used to aid the information-seeking process. Whilst search personalisation approaches go by a different name they share much in common with recommender systems, as both attempt to identify subsets of relevant items on the user's behalf. The next section will examine the two major classes of recommender systems, one of which takes an explicitly collaborative approach.

2.4. *Information Filtering with 'Classic' Recommender Systems*

Recommender systems aim to help users identify items that might be relevant to their needs, and are commonly used for tasks such as *suggesting* related items to users of an e-commerce Web site (Schafer, Konstan and Riedl, 2001), or *filtering* a set of documents such as emails or newsgroup postings (Goldberg, Nichols, Oki et al., 1992; Hill and Terveen, 1996; Terveen, Hill, Amento et al., 1997) to exclude those that are less relevant.

Whilst the recommender systems domain is sometimes conceived as solely concerned with research into the former of these functions, it should be noted that the *suggesting* and *filtering* tasks are isomorphic (Belkin, 2000). Both involve using system input to reduce a larger set of initial items of low average relevance to the user, to a smaller set of more relevant items, where the nature of the relevance is determined by and encoded in the recommendation algorithm.

Consequently, as has been noted by authors such as Ansari, Essegaiier and Kohli (2000), Celma (2006) and Schafer, Konstan et al. (2001), both e-commerce recommenders and Web search engines function as recommender systems; search engines simply

recommend items from their corpus of crawled Web documents, based on user keyword input. The primary difference between these two deployment contexts is therefore merely conceptual, based on who initiates the recommendation process (the system or the user) and on the perceived goal of providing recommendations (encouraging the user to buy more products vs. finding relevant documents).

Input to recommender systems may be provided explicitly by the user or generated automatically by their interactions with the system, and may take a number of forms: a user profile (such as interests or demographics), a seed item the user has already interacted with in some way (e.g. past purchases from an online shop), or keyword terms related to the user's information need.

Perhaps the most rigid approach to creation of profiles may be to ask users registering with a site to specify topics in which they have an interest. This would be time-consuming, requiring the user to map their interests to a third party topic hierarchy with which they may be unfamiliar, and inaccurate due to a granularity mismatch in listed topics. Furthermore, in a domain-specific system the user may be required to undertake this process when they have just one specific and short-lived information need, making the cost-benefit ratio high. By contrast, in a generic system the user may have to specify a full range of interests in advance, trying to anticipate future information needs they may have for which creating the profile might be important. For these reasons, systems that take implicitly generated input or minimise the upfront investment required by the user are generally preferred.

2.4.1. Content-Based Recommendation

Recommender systems generally follow one of two approaches. *content-based recommendation* or *collaborative filtering* (Balabanovic and Shoham, 1997). Content-based recommendations can be made in a number of ways: by matching the content of an item to some input such as a user profile (e.g. Balabanovic and Shoham, 1997) or keyword terms (in the case of a search engine); or by matching the content of an item to that of another item for which the user has already shown some preference. For example, action films may be recommended to the user where they have expressed an interest in this type of film, or satirical comedies where they have previously viewed or purchased items from this genre. Web search engines are an example of the content-based method, whereby results are returned based on content matches between documents and user-supplied keywords. Content-based recommendation is not limited to textual documents, and can be applied to other media formats. For example, Celma (2006) has used the approach to recommend musical artists based on the characteristics of their music.

2.4.2. Collaborative Filtering

In a contrasting approach, collaborative filtering recommender systems need know nothing of the content of items they recommend, relying instead on the actions of others to identify the most appropriate items. As such, collaborative filtering systems provide rudimentary support for word-of-mouth recommendation (Shardanand and Maes, 1995). However, this support is not particularly sophisticated, as the individuals on whose profiles recommendations are based are only linked to the user statistically.

The kinds of behavioural traces or actions on which collaborative filtering recommendations can be based does vary somewhat. The label 'collaborative filtering'

was initially introduced by Goldberg, Nichols et al. (1992) to describe their *Tapestry* system for filtering and prioritising incoming text documents by reusing filters defined by others using the system. In this initial form, the collaborative aspects of collaborative filtering were rather rudimentary (in contrast, the filtering capabilities available through the *Tapestry Query Language* were fairly sophisticated), and involved knowing the names of others who created useful filters, the names of the filters themselves, and the task context in which they were useful. Whilst cumbersome, this approach does have the benefit of allowing the user to manually select trusted sources for filters based on their knowledge of the person.

Moving away from the reliance on known colleagues to provide reusable filters, Hill and Terveen (1996) and Terveen, Hill et al. (1997) describe PHOAKS, a system for recommending Web pages on certain topics by mining postings to newsgroups. The system benefits from reusing the inputs of large numbers of users but has rather basic ranking mechanisms (based on frequency of mention of Web links) and a rather weak notion of recommendation, whereby including a link in a newsgroup posting (excluding certain cases such as links in signatures) constitutes a recommendation. In Terveen, Hill et al. (1997) the authors claim that text surrounding links is analysed to look for 'markers' that indicate a recommendation, but how this process operates or which words constitute a marker is not elaborated. Having ignored the identity of the contributor when computing recommended links, the PHOAKS system does display this information alongside results, thereby providing a means for interested users to find out more about or contact contributors. This suggests that the authors see value in being able to form richer impressions of the source of a recommendation, presumably as a means to assess the quality or relevance of that recommendation.

Subsequent work in the field has rather commandeered the collaborative filtering label to mean systems that recommend items by correlating a profile of the user with that of other unknown users of the system, and recommend items for which the correlated users have indicated a preference (Herlocker, Konstan and Riedl, 2000). This approach substitutes more subtle judgements based on knowledge of the source for recommendations based on larger statistical trends.

To illustrate this form of collaborative filtering with an example, two users A and B may have each purchased a number of items from the same online store. In doing so they have each implicitly expressed a degree of preference for these items. Where there is a high degree of overlap between the items bought by A and B (they share a 'co-preference' for a number of items) these two users are presumed to share similar tastes. Therefore if A purchases an additional item, there is deemed to be a high probability that B will also be interested in that item. The degree of correlation between users is often referred to as 'taste overlap', and serves as a predictor of the accuracy of the recommendations.

In addition to binary data about which items a user has previously purchased (or 'consumed' in some other way), user profiles may be based on explicit ratings (binary or numerical) given by the user to items in the system. Some deployments of such rating functionality are described in Josang, Ismail and Boyd (2007).

This correlational, person-to-person approach to collaborative filtering is widely used in so-called *taste domains* (Bonhard, Harries, McCarthy et al., 2006), where the heterogeneity in product choices and preferences is accounted for primarily by differences in consumer tastes (Ansari, Essegai et al., 2000). Collaborative filtering systems have been deployed in taste domains such as music (Shardanand and Maes,

1995), Usenet postings (Konstan, Miller, Maltz et al., 1997), and films (Miller, Albert, Lam et al., 2003).

The socially-oriented music site *last.fm*² recommends music by mining listening habits to identify taste overlaps between users. This approach arguably creates more sensitive measures of similarity between users than those based on manual trust ratings or purchasing behaviour; repeated listening suggests ongoing taste for a particular item whereas a one-time purchase may have been made in error.

Irrespective of their accuracy, however, these measures of similarity between individuals are not global, but music-specific. Therefore they cannot necessarily be assumed to indicate similarity between individuals across heterogeneous domains, or in domains from which they were not originally derived: "*...agreement in one domain...is not necessarily predictive of agreement in a different domain...*" (Konstan, Miller et al., 1997, pp. 83).

Among a number of challenges for the next stage of research into recommender systems, van Setten, McNee and Konstan (2005) recognise the need for cross-domain recommendations, and the limitations of existing systems in supporting these.

While it is likely that taste to some extent permeates human judgements in all domains, the degree to which it influences decision-making processes is likely to vary according to the characteristics of the domain, and the specific demands each domain places. For example, while the decision to buy a low-cost book may be mediated primarily by one's

² <http://last.fm/>

taste in fiction, the decision to accept a particular stock market investment or medical treatment may be more influenced by one's financial resources or the level of risk involved.

While even factors such as aversion to risk arguably reflect an aspect of one's general tastes, there likely remains a spectrum of recommendation domains along which the influence of taste varies. One end of this spectrum may be occupied by domains (such as music, or recipes (Konstan, Miller et al., 1997)) where decisions are heavily mediated by taste. The other end may be occupied by domains (e.g. household appliances) in which objective criteria such as functionality and practical considerations are of greater importance.

The underlying question, however, is not simply whether taste correlates across domains situated along this spectrum, but whether they correlate sufficiently to support collaborative filtering and whether adequate metrics can be derived to capture these relationships (Dieberger, Dourish, Höök et al., 2000). For example, whilst one's preferences for music may be predictive of one's taste in films, it is not readily apparent how well taste in either of these domains can predict the household appliances one chooses.

This raises the question of whether personalisation and recommender systems in their current form are limited to operating purely in more taste-centric domains, where preference for one item can be highly predictive of preference for another.

In domains where taste is just one of many important factors, there may be a greater role for recommendations from those with significant expertise in or experience of the domain (Dieberger, Dourish et al., 2000), irrespective of the similarity in taste. Accepting

recommendations from such sources may require more complex or in-depth judgements about their trustworthiness or relevance. This may in turn influence the data and techniques that can be used to support the recommendation-seeking process, and highlight a need to explore novel recommendation approaches.

One variation in approach to collaborative filtering is described by Linden, Smith and York (2003) who, for reasons of scalability, use correlations in the purchase profile of items as the basis for recommendations, resulting in item-to-item rather than person-to-person collaborative filtering. In both cases, strong correlations between users or items form the basis for providing recommendations.

A number of attempts have also been made to create hybrid recommender systems using both content-based and collaborative filtering approaches. For example, Balabanovic and Shoham (1997) use a hybrid approach to recommend Web pages, whilst Salter and Antonopoulos (2006) apply similar principles to films.

Authors such as Resnick and Varian (1997) have questioned the label 'collaborative filtering' as, particularly in systems that follow the more recent usage of the term, users may not be collaborating at all in the formal sense of the word. In contrast, the users of most systems are completely unknown to each other (at least the identity of other users is not apparent from the output of the recommender system). Consequently, collaborative filtering recommender systems do not currently reflect how people seek information using social networks of people they know.

Bonhard and colleagues (Bonhard and Sasse, 2005) argue that recommender systems research to date has largely ignored the social context of recommendation-seeking. when in fact this may provide many benefits. Integrating social networks with recommender

systems may help to generate recommendations that are more useful, trustworthy, and comprehensible, thereby lowering the cognitive effort for the user in judging their relevance. A number of approaches that fall under the umbrella of social navigation provide greater opportunities for collaborating or interacting with known or pseudonymous individuals during the information-seeking process.

2.5. Social Navigation

Social navigation is a design approach that aims to utilise the presence and actions of people in online environments as a means to assist others in navigating the same virtual spaces. Therefore users may be supported in locating and evaluating information and subsequent decision making, through mechanisms such as visualised traces of other peoples' activities, or direct communication channels (e.g. chat) with other users of a system (Dieberger, Höök, Svensson et al., 2001, Dieberger, Dourish et al., 2000).

The term was originally coined by Dourish and Chalmers (1994) in order to distinguish between social navigation (based on information from other people) and semantic navigation (based on the underlying structure of the information being navigated).

Whilst recommender systems, and in particular collaborative filtering applications that reuse the efforts or actions of other people to filter information, have been treated as one form of social navigation, there are many other avenues of research that fall under the same label. In fact, the nature of social navigation has been interpreted fairly broadly, giving rise to a wide range of applications and approaches. For example, the *Footprints* system (Wexelblat and Maes, 1999) uses visio-spatial metaphors such as maps and paths to indicate how previous users interacted with a Web site, whilst Svensson, Hook,

Laaksolahti et al. (2001) bring together a number of social navigation features such as chat, recommendation, and avatars in a system for navigating food recipes.

Mobasher, Cooley and Srivastava (2000) describe a proof of concept system, called *WebPersonalizer*, that suggests potentially relevant pages to the user while they browse the site. Whilst not explicitly labelled as an example of social navigation, this application follows the same principles. Suggestions are made based on data about how previous users have navigated the site, obtained by analysing Web server logs. The analysis is performed anonymously therefore all users who follow the same navigation path on the site receive the same suggestions. This has the potential to be rather self-reinforcing, whereby all users are channelled along similar paths irrespective of their underlying information need or task.

In addition to social support for browsing, *social search* has been investigated. The term 'social search' can be interpreted in two ways. The first of these (referred to as 'Social Search (Items)' in Figure 1) falls under the umbrella of social navigation and sees social search as supporting conventional search processes with information derived from the actions or preferences or other people.

This first interpretation of the term is adopted by researchers such as Ahn, Brusilovsky and Farzan (2005) who explore the use of page visit data and user annotations to supplement search results in their *Knowledge Sea* application. Search results are ranked using a conventional document ranking technique (in this case *TF-IDF*) and then supplemented by displaying users' own visit frequency for particular documents alongside aggregate visit data from a wider group and indications of the degree of 'praise' the document has received. The use of page view data and endorsements (in the form of

positive or negative praise) in the results interface bears many similarities to the use of customer purchase data or ratings in collaborative filtering recommender systems. However, the nature of the group from which aggregate statistics are drawn is not specified, and as with collaborative filtering performance is reliant on the behaviour of other anonymous users.

The second interpretation of the 'social search' label ('Social Search (Network)' in Figure 1) refers to searching a social network to identify particular individuals who may be able to assist with the current task. Work that follows this definition is discussed in Section 2.6.1 below.

The importance of maintaining privacy in social navigation systems has been raised (e.g. Dieberger, Höök et al., 2001). However, it is also argued in the same paper that a degree of *visibility* is essential in order for applications to retain utility, which certainly points towards pseudonymous and possibly even towards known identities in social navigation systems.

In support of the arguments put forward by Bonhard and colleagues described above, and counting against anonymous applications, Kautz, Selman et al. (1997b, 1997a) argue that not all information sources are equally desirable. Consequently, personal referrals between known individuals allow the information seeker to make judgements about the quality of the information they are receiving and may instil greater confidence in the information if the source is trusted.

How people select sources for information and recommendations will be reviewed in detail in Chapter 4. However, before this, research and systems will be discussed that

attempt to integrate more directly with known social networks, and utilise these to support the information-seeking process.

2.6. *Approaches Based on Social Networks*

2.6.1. Social Search (Network) and Expert Finding

In contrast to the first (item-oriented) interpretation of social search as followed by e.g. Ahn, Brusilovsky et al. (2005), others relate the term to the process of searching one's social network for a particular piece of information, or a person who has access to that information. Therefore, in this case the social network is not simply a source of data that can be used to aid one's search, but the source itself; the search task becomes a search for the appropriate node in the network.

Kautz, Selman et al. (1997a) argue that "many information-gathering tasks are better handled by finding a referral to a human expert rather than by simply interacting with online information sources" (pp.27). They describe a system *Referral Web* which aims to addresses exactly this issue by data mining the Web to build models of social networks connecting researchers, and to identify the areas in which each has expertise. Social networks are constructed by identifying co-occurrence of names in Web pages, whilst person-topic relationships are inferred from name-topic co-occurrence on Web pages, where topics are taken to be "capitalized phrases that appeared in documents...but were not proper names" (pp.33).

Having built these models, the application then allows a user to view members of their extended social network who have expertise in an arbitrary topic, or to view a path between themselves and a particular individual, even where neither individual has

actively registered with the system. This demonstrates the value of priming systems with background data from the Web.

Searching for those with expertise relevant to problem-solving tasks in workplace or organisation settings has also been extensively investigated by McDonald and Ackerman (2000). Based on the findings of an earlier field study (McDonald and Ackerman, 1998), they present a generic 'Expertise Recommender' architecture and a specific implementation of this architecture tailored to one organisation. McDonald and Ackerman argue that expertise seeking methods are heavily embedded in local settings and work practices, and instantiate their system at a correspondingly specific level. Whilst these arguments would seem to make intuitive sense and have the support of their earlier research, the resulting systems would appear to require considerable customisation to be useful in any one setting. Furthermore, it is not apparent how applicable these findings (and the corresponding architecture) would be outside of a workplace setting. The authors do not discuss other contexts in which expertise recommendation may be required, attempt to deploy systems in other settings, or explore its potential utility in domains where expertise is of lesser importance.

2.6.2. Social Networks and Trust

A number of attempts have been made to enhance social network-based approaches to information-seeking with notions of trust. In most cases trust is employed as a fairly broad, non-specific concept. These attempts are examined below, and can be analysed according to four dimensions:

1. **automation**: the degree to which trust ratings are automatically computed (versus provided manually)

2. **topicality**: the degree to which trust ratings are topical in nature (versus one global trust rating of an individual across all topics)
3. **individuality**: the degree to which trust ratings are personal (versus one global trust rating of an individual shared by all others)
4. **anonymity**: the degree to which the system operates over networks of known individuals (versus operating across systems of unknown individuals, or a mixture of the two)

Golbeck and Hendler reach beyond the network of personally known individuals by combining social networks and inferred trust/reputation relationships in an email filtering application (*TrustMail*) (Golbeck and Hendler, 2004) and film recommender system (*FilmTrust*) (Golbeck and Hendler, 2006).

The goal of *FilmTrust* is not to actively suggest items to the user unprompted, but to provide her with feedback on how likely she is to be interested in a film she has already found, based on direct or inferred trust relationships. Film reviews are also ranked on the same basis when displayed on the site. In a similar fashion, *TrustMail* annotates each email in the user's inbox with a trust rating, based on trust relationships computed through the network between sender and receiver.

To participate in the trust networks associated with these applications, and benefit from their filtering and ranking capabilities, the user must manually rate (on a 1-10 scale) the reputation of, or their trust in, people they know. In *TrustMail* these ratings are non-domain-specific 'reputation' ratings of the known person, whereas in *FilmTrust* the user is asked to rate her trust in the person in the context of films. These ratings then seed the algorithmic creation of trust scores for all other members of the wider network to whom

the user is linked socially. Importantly, these scores are computed from the user's local perspective, rather than being global to the entire network. This work is characterised by a mixed approach to *automation*, no *topicality* in the *TrustMail* system but a limited amount in *FilmTrust*, a high level of *individuality*, and varying degrees of *anonymity*.

The approach is useful in that it enables trust ratings to be inferred between individuals who are connected to some degree, but do not know each other personally. This can be of value where insufficient information is available within one's immediate network, or one's immediate network is too small. In addition, there is some evidence (Golbeck and Hendler, 2006) to suggest that this approach can produce more accurate results than 'nearest neighbour' collaborative filtering techniques in situations where the user's tastes are divergent from the population as a whole.

However, the approach has a number of limitations. Firstly, the semantics of the trust relationships are often ambiguous or underspecified. In *TrustMail* users are asked to rate the general reputation of people they know. Whilst reputation may not be quite so context dependent as trust, this still appears to be a gross oversimplification. For example, a researcher may have an excellent academic reputation, but be known to be unreliable when repaying loans. In the context of email filtering the risks associated with this are small, however under-specifying relationships in this way does limit the value and reusability of the data.

The ontology Golbeck, Parsia and Hendler (2003) used to describe the trust ratings provided by users in *TrustMail* and *FilmTrust* does in fact allow specification of the topic or domain in which the trust is being asserted, and whilst users of the *FilmTrust* system

are asked to provide trust ratings in the context of film reviews, this relationship is not explicitly encoded in output from the system.

Secondly, this approach does rely on provision of manual trust ratings between users to bootstrap the process. Whilst making just one social connection in the *FilmTrust* network does allow recommendations to be made for a user for 95% of films, it would be desirable to investigate existing sources of information from which trust relationships between known individuals could be inferred, in order to bypass this manual annotation process and lower the cost of participation for users.

Thirdly, the work of Golbeck and colleagues uses trust ratings as the basis for making similarity assessments between users. This is justified by reference to work by Ziegler and Lausen (2004) that found a correlation between trust and user similarity in the online community *All Consuming*. Whilst trust may serve as a valid proxy for similarity, this correlation may be due to a third factor which has not been accounted for, and as such the predictive validity of this relationship should be questioned.

Numerous other attempts have been made to integrate notions of trust with social networks. For example, Richardson, Agrawal and Domingos (2003) describe a distributed 'web of trust' approach, intended to support the assessment of 'belief' in assertions on the Semantic Web as a function of the user's subjective trust in the author of the statements. The approach assumes that no one entity will know the trustworthiness of every other, and therefore ratings cannot be assigned to an entity by a central source. On this basis, the authors propose that each user specifies a set of other trusted users, and a recursive propagation model is then used to compute a user's trust in all other connected members of the trust graph. This results in moderately *automated* trust ratings that are

individual in nature, and therefore trust in the same entity may vary significantly between users. This user-centric model of trust is compatible with the perspective taken in this dissertation, as it gives a more personal view of the network. The approach of Richardson et al. does not support the specification of trust *topicality*, although this is raised by the authors as an area for future research.

Due to their statistical foundations, collaborative filtering systems require data sets of a significant size in order to perform at optimum levels. Massa and colleagues (Massa and Bhattacharjee, 2004, Massa and Avesani, 2004) use review and web of trust data from the reviewing site *Epinions*³ to demonstrate how trust propagation techniques can be used to overcome the *cold-start/early-rater* and *sparsity* problems that can affect conventional collaborative filtering approaches.

The *cold-start* problem refers to situations in which items added to the catalogue of an e-commerce Web site can not be recommended using collaborative filtering until at least one customer purchases that item. Only at this point (and assuming that the customer already has purchases in common with other customers) can predictions be made of which other customers may be interested in the item. The extreme cold-start situation is that of a totally new recommender system where no data exists with which to correlate users or items.

Cold-start affects users in a similar fashion, as they must develop a profile that correlates them with other users (perhaps by rating or purchasing some items) before recommendations can be provided (Massa and Bhattacharjee, 2004). Early-rater

³ <http://www.epinions.com/>

problems (Dieberger, Dourish et al., 2000) describe one specific aspect of this situation, in which early adopters of a system gain little performance benefit in return for their input, as the system as a whole is not sufficiently populated with comparable users on which to base recommendations.

Sparsity is a measure of the degree to which items or users in a collaborative filtering system can be compared. Systems where users can on average be compared to a relatively low number of other users (due to a lack of overlap in profiles) are described as 'sparse', and will tend to provide lower quality recommendations (Massa and Avesani, 2004). These factors can all limit the ability to deploy recommender systems in settings where only small data sets are available on which to base recommendations.

Existing data from external sources is not commonly used to help bootstrap recommender systems. This is likely due to a lack of relevant data being available in an easily reusable form, either from the Web at large or from existing recommender systems. Issues such as privacy, data protection and maintaining competitive advantage reduce the incentives to share profile data, leading to duplication of effort by users who cannot benefit from using aggregate profiles of their own data across multiple recommender systems. If more data (such as reviews or broader profile information) were to be published online in an easily reusable form, this may provide a source of background data with which to bootstrap recommender systems, thereby reducing cold-start and sparsity issues.

Massa and colleagues (Massa and Bhattacharjee, 2004, Massa and Avesani, 2004) show that propagating trust through the network as a function of inverse network distance can provide systems with greater coverage of users and items on which to base

recommendations, whilst keeping error relatively low. This is particularly useful when providing recommendations to new users who have not rated many items. Whilst these findings suggest there may be a role for this form of trust propagation, more sensitive trust metrics are required as the simplicity of the trust data on which it is based may be a limiting factor. See 2.7.6 for a full discussion of this issue.

In relation to the work of Golbeck and Hendler (but equally applicable to the related studies discussed above), O'Hara, Alani, Kalfoglou et al. (2004) observe that trust is not strictly transitive, and highlight this as a potential shortcoming of the work. This criticism applies to all the approaches described above that use trust propagation in order to compute metrics for indirectly connected (and therefore unknown) members of a social network. The results obtained by Golbeck and Hendler (2006) comparing their approach to collaborative filtering suggests that this may not significantly reduce the utility of the system in the context of film reviews. However, it may be that in domains less mediated by taste and where greater risk is involved, simple trust relationships such as these may not be so reliably propagated through an unknown network.

2.7. Conclusions and Gap Analysis

The research reviewed in this chapter suggests that collaborative and social processes have a powerful role to play in reducing information overload and increasing personal relevance in information-seeking, through filtering, recommending and ranking. Substantial work has been carried out in these areas, producing techniques and systems that are now in widespread use (Schafer, Konstan et al., 2001). Despite the benefits afforded by these approaches, a number of gaps are present in existing work.

2.7.1. The Nature of Relevance

The literature reviewed above highlights that while relevance in information-seeking is often treated as a global, topical relationship between a query and a set of items, there is a strong case for viewing it as a more subjective relationship between items and the user's abstract information need. This raises the question of how to capture or represent underlying information needs, particularly in the light of work by Teevan, Dumais et al. (2005) suggesting that these may not be easily and unambiguously expressed as keyword searches.

Existing approaches such as personalised search sidestep this issue, and attempt to identify relevant items by interpreting input (e.g. search terms) in the light of a profile of the user. In doing so they are implicitly using background information about the user to predict which items may be relevant, thereby implicitly inferring his or her information needs at a very general, non-specific level. This raises the issues of which forms of profile information are most predictive of relevance and underlying information needs.

Furthermore, what constitutes relevance is likely to vary significantly according to the demands of the task and the context in which it occurs. As a result, the model and data used to determine relevance should also vary, and this must be taken into account by systems which aim to support a wide range of tasks and task contexts. Current systems do not do this, as they tend to be neutral with respect to the task.

2.7.2. Economies of Scale through Collaboration

In many cases, personalised search and content-based recommendation are limited by their failure to capitalise on economies of scale through collaboration with other users of

a system. The only source of profile information from which more personal relevance can be determined is the user himself, who must invest heavily in building his own profile in order to benefit from using the system; opportunities to 'piggy-back' on the knowledge or actions of others do not exist.

Social navigation systems, in particular collaborative filtering recommender systems, aim to address this shortcoming, by using the behavioural patterns and preferences of others to support an individual's information-seeking process. Profile information about other people can be combined with user input to help identify relevant items, but only as long as other people can be identified whose profiles are in some way relevant to the user's information need.

Collaborative filtering recommenders exploit taste overlap as a proxy measure of the likely relevance of one user's profile information to the needs of another user; if two users A and B share a significant taste overlap, there is a reasonable chance that information about the preferences of A can be used to help identify items relevant to the needs of B , but *only* in domains heavily mediated by personal taste.

2.7.3. Personalisation and Recommendation across Varied Domains

Identifying relevant profiles outside taste domains remains a major challenge that is poorly supported by current systems; collaborative filtering is not adapted to situations where the user requires recommendations from a domain expert, irrespective of their taste overlap.

Expert finding systems such as *Referral Web* (Kautz, Selman et al., 1997b, 1997a) take a source-centric rather than item-centric approach to information-seeking, whereby identifying the most appropriate human source of the information is the key search challenge in order to obtain relevant information. This can be seen as another instance of the process mentioned above of matching information needs to the profiles of other people; however, in this case the challenge is to match expertise profiles to information needs rather than taste profiles to taste profiles.

What is lacking from existing systems is the flexibility to provide personally relevant information or recommendations across a wide range of domains and tasks. This shortcoming may have both technical and theoretical underpinnings: systems may be algorithm-centric, where use of a particular technical approach defines the functionality available in the system, rather than need-centric, where an identified user need determines the functionality of a system and thus its underlying technical approach. Contributing to this may be a lack of theoretical understanding of how information needs vary across domains, tasks and contexts.

2.7.4. Personalisation and Recommendation in Open Worlds

The personalised search and recommender system approaches discussed above all operate on relatively fixed, predefined sets of items or users – 'closed worlds'. These closed worlds may be small, limited to jobs listed on a particular job-seeking site (e.g. Bradley, Rafter et al., 2000), or larger, such as all items in an e-commerce store, but they are closed nonetheless. Consequently, only items that are represented in the system can be recommended to users or presented in search results. If the user need cannot be met by any items in the system then performance is reduced, as potentially suitable items remain

outside the closed world. Unfortunately, existing systems are generally poor at suggesting alternative sources of recommendations or personalised search results that may be more appropriate to the user's information needs.

Despite their size, and irrespective of their degree of coverage of the Web, the indices of the major search engines (e.g. Jeh and Widom, 2003; Teevan, Dumais et al., 2005) also represent closed worlds. Much of the information that people might seek is not available online and probably never will be, either because it is personal or sensitive in nature; is stored in a legacy format or system (Kautz, Selman et al., 1997b, 1997a); or is simply too complex to represent in a computational system. Consequently, the space of human knowledge and the space of results that can be represented in an information-seeking system are unlikely to ever fully overlap.

By this definition, resolving the closed world issue in technical systems is not feasible. However, what should be investigated are approaches that can broaden the scope of knowledge represented in search and recommender systems, particularly to encompass legacy information that is not available online, or information that is too complex to be represented in a computational system.

2.7.5. Sparsity and Cold-Start Problems

Systems large and small can suffer from sparsity and cold-start problems, due to the cost of bootstrapping the system with sufficient data. The literature reviewed above suggests that in such situations social networks can provide a viable basis for approaches that can reduce these issues. For example, trust propagation has been used to mitigate problems caused by sparsity, whilst maintaining the accuracy of recommendations (e.g. Golbeck

and Hendler, 2004, Golbeck and Hendler, 2006, Richardson, Agrawal et al., 2003, Massa and Avesani, 2004).

However, these applications simply use social network connections as a conduit to other, unknown parts of the network in order to compensate for sparsity. This approach is predicated on two assumptions: firstly, the principle of *homophily*, i.e. that people are more similar (across a number of dimensions) to members of their social networks than to members of the population at large (McPherson, Smith-Lovin and Cook, 2001); secondly, that similarity equates, or at least correlates, with trust, as found by Ziegler and Lausen (2004). The first of these assumptions is widely accepted. The second, however, should not be taken as robust outside taste domains without further research; situations that require extensive domain knowledge may warrant recommendations from individuals who are highly dissimilar to the information seeker, but highly knowledgeable in a relevant domain.

Literature in this area will be discussed in more detail in Chapter 4. However, in the meantime it should be noted that overcoming sparsity and cold-start issues in non-taste domains requires an approach that goes beyond the social network trust propagation methods discussed above.

2.7.6. Richness of Trust Models

Collaborative filtering recommender systems and related applications make recommendations or relevance assessments by correlating a profile of the user with those of other unknown individuals. While this maintains privacy and therefore enables large data sets to be used, it prevents users from applying their own knowledge of information sources when judging the trustworthiness of recommendations. Just as the algorithms

used in recommender systems are often seen as 'black boxes' (Herlocker, Konstan et al., 2000), so may the users who generated the data on which recommendations are based.

In addition, the data on which transitive trust relationships are calculated is often of relatively poor quality. In cases such as *Epinions*, the relationships in the web of trust are very coarse, being binary in nature (Massa and Bhattacharjee, 2004) and with rather ambiguous semantics. There is no requirement that two parties are known to each other; one user may add another to their web of trust simply based on having read some of his reviews.

Such relatively unsophisticated trust models do not adequately reflect how people seek information and recommendations from those around them across a wide range of tasks and contexts, or the mechanisms they use to infer the trustworthiness of these sources. Consequently, applications that implement these approaches are not well adapted to supporting a broad range information-seeking scenarios and compare unfavourably with information-seeking based on word-of-mouth recommendation through traditional channels.

With few exceptions (e.g. Kautz, Selman et al., 1997b, 1997a), systems make little direct use of known members of one's social network with whom one has *existing, nuanced* relationships and about whom one can make detailed, considered trust judgements. There is clear room for new approaches in producing social filtering and recommendation applications that truly capitalise on these characteristics.

Developments in this area require a greater knowledge of the information- and recommendation-seeking process among people who know each other personally, and the factors that guide these decisions. Bonhard and Sasse (2005) make progress in this

direction, but their work remains constrained in taste-mediated domains. This kind of improved understanding may yield technical approaches and systems that are not constrained to supporting limited tasks or contexts.

2.7.7. Summary

In summary, systems are needed that support a more personalised notion of relevance in information-seeking. Such systems must take into account that what constitutes relevance is likely to vary according to the characteristics of the information-seeking task; consequently, they should be designed to operate across a wide range of contexts, not simply in taste domains *or* expert finding.

Meeting these requirements may involve systems being more grounded in conventional information-seeking approaches such as word-of-mouth recommendation, and more supportive of users applying their own knowledge to assess the trustworthiness of recommendation sources. This may in turn require a deeper theoretical understanding of how people choose information and recommendation sources, the factors that guide these decisions, and how these vary across tasks with different characteristics in order to ensure relevance.

The outcome of such an investigation is likely to affect the data requirements of systems that adopt this approach. Sparsity and cold-start issues may be equally problematic, and possibly more so if a wider range of information-seeking contexts are being supported. Mechanisms for overcoming these issues should be investigated.

Lastly an approach is needed that can reduce the extent to which search and recommender systems represent closed worlds, and open these up to including information that is not otherwise available online.

The following chapter aims to address these limitations through a social network-oriented approach to Web-based information-seeking.

3. Approach: Personalised Relevance in Information-seeking through a Trusted Social Network

The Web has indisputably demonstrated its capabilities as an information sharing and dissemination platform. However, from the gap analysis in the previous chapter it is apparent that information-seeking applications on the Web would benefit from:

- adopting more personalised notions of relevance
- supporting a wider range of information-seeking tasks, which may vary in their characteristics
- being sensitive to how variations in task characteristics may determine relevance
- enabling greater involvement of the user's own knowledge in the information-seeking process
- broadening their scope to include information that may not be available online

Social networks have long provided a powerful means for obtaining relevant and trustworthy information. This research proposes to address the shortcomings listed above by exploiting synergies between the Web and social networks. The outcome of the research will be approaches and systems that support information-seeking on the Web by harnessing the knowledge and experience of the user's social network, according to the principles of word-of-mouth recommendation. The aim is to increase personal relevance

and facilitate greater use of trust, thereby improving the effectiveness of information-seeking and reducing information overload.

Numerous previous attempts have been made to support word-of-mouth in a Web environment, through, for example, collaborative filtering and online reviews. This research is not intended to replace these, but instead to develop complementary approaches and technologies that can overcome identified limitations in existing work. The factors outlined below distinguish this approach from previous work in the area.

3.1. Characteristics of this Approach

3.1.1. Source-centricity

In contrast to many of the search and recommendation approaches discussed in Chapter 2, this research takes a source-centric rather than item-centric approach to the information-seeking process; i.e. the emphasis is on identifying relevant sources before trying to identify relevant items.

The first challenge of this approach is *source identification*: finding out whom within a social network *knows* about topics relevant to the information need and therefore may be able to provide relevant information or recommendations. The second challenge is *source selection*: deciding which of these individuals to *trust* as sources of personally relevant information and recommendations. This research aims to develop approaches and systems that address both these challenges.

The reader may be interested to note that source identification and source selection can be seen as generalisations of McDonald and Ackerman's (1998) *expertise identification* and *expertise selection* discussed in more detail in Chapter 4. Regarding the concept of

trust, many definitions have been proposed in the literature, of which Marsh (1994) provides a thorough review. For the purposes of this dissertation, and in the context of word-of-mouth recommendation-seeking, trust is defined here as '*confidence in another individual as a source of accurate and relevant information*'. This definition is deliberately neutral with respect to the source of evidence on which this confidence may be based.

3.1.2. Task-adaptivity

By definition, any information-seeking process must be aligned to the demands of the task by which it was originally motivated. This task will not only define the information need, but is also likely to have a number of other characteristics that will determine what constitutes an appropriate source of information or recommendations. This research aims to further understand these characteristics, and develop source identification and source selection processes that are sensitive and adaptive to them.

3.1.3. Social Networks and this Approach

The role of social networks in online environments, and online environments as reflections of social networks themselves, has received increasing attention in recent years. Garton, Haythornthwaite and Wellman (1997) emphasise the value of a social network perspective in the study of computer-mediated communication, and summarise some of the key units of analysis in the field of *social network analysis* (Scott, 2000).

Of particular relevance to the research presented here are the notions of *relations*, *ties* and *ego-centricity*. One or more *relations*, such as sharing information or being members of the same organisation, create a *tie* (often classified as weak or strong) that connects a

pair of actors. Research into the roles of strong and weak tie relationships is discussed in more detail in Chapter 4.

Garton et al. distinguish between *ego-centric* or *whole network* views of social networks. The ego-centric approach views the network from the perspective of a particular individual, whereas the whole network view considers an entire network comprised of individuals who meet a certain criterion. The former, ego-centric perspective on social networks is of greater relevance to this research.

Authors such as Mika (2004) have studied how information available on the Web reflects the structure of social networks in the offline world. By combining data harvested from the Semantic Web with conventional Web mining approaches, he is able to identify structural properties of the social network within the Semantic Web research community, such as various measures of each member's *centrality* within the network.

These metrics provide a basis for understanding some of the structural properties of a particular social network. As the research reported here is concerned primarily with the nature of one-to-one relationships in social networks, and the implications of these for information- and recommendation seeking, these measures of the structural attributes of social networks will not be considered in further detail.

In addition to ongoing work examining social networks themselves, whether online or offline, there has been an increasing interest in developing Web applications that include a social component. For example, the primary emphasis of sites such as *Facebook*⁴ and

⁴ <http://www.facebook.com/>

*LinkedIn*⁵ is in allowing people to express the connections in their social networks, forge new connections and engage in social interactions online.

In slight contrast, social annotation and bookmarking services, such as those summarised by Hammond, Hannay, Lund et al. (2005), allow individuals to store and annotate items for their own usage, but also share these resources with others through the social networking aspects of the sites.

Current trends in Web applications and ongoing research into social networks increase our understanding of the interaction between social factors and online environments, and provide a context for the research presented here. However, rather than looking at social networks purely from a structural/analytical viewpoint or from the perspective of technical applications, the research presented here requires a fuller understanding of how information and recommendations are sought within social networks, and the factors that shape this process. These will be examined in detail in Chapter 4.

3.2. Benefits of this Approach

3.2.1. Increased Personal Relevance

One fundamental premise of this approach is that members of one's social network are more likely to have knowledge relevant to one's own information needs than are people outside one's network.

⁵ <http://www.linkedin.com/>

The homophily principle (McPherson, Smith-Lovin et al., 2001) mentioned in Chapter 2 states that people are more similar (across a number of dimensions) to members of their social networks than to members of the population at large. Whilst any member of the population may have knowledge relevant to a particular information need, the increased similarity stemming from homophily suggests that the knowledge held by members of one's social network will be of greater personal relevance. The approach pursued in this research aims to exploit this characteristic to support information-seeking.

In a related but not equivalent fashion, collaborative filtering exploits a relationship between taste overlap and relevance. However, a point to be noted is that similarity and taste overlap are not being equated in this research. Whilst the two are likely to correlate to some extent, similarity from the perspective of homophily encompasses many more dimensions than simply taste overlap, and should therefore be seen as a broader concept.

What is being proposed here is a positive relationship between personal similarity and perceived relevance; the greater the similarity between two individuals across a number of dimensions, the greater the likelihood that they will find the same information or items relevant to their information needs. By focusing on the relationship between similarity and relevance (rather than taste overlap and relevance) the approach taken in this research aims to be applicable beyond simply taste domains.

It should also be noted here that similarity is not being equated or correlated with trust, as in the work of Ziegler and Lausen (2004). Research presented in later chapters of this dissertation demonstrates that trust is a nuanced and task-dependent concept that may only correlate with similarity under certain circumstances.

3.2.2. Utility across a Range of Tasks

While similarity may provide a sound basis for increased personal relevance, the strength of this relationship is likely to vary according to the characteristics of the task that motivates the information-seeking, in which case additional factors will need to be taken into account in determining the relevance and trustworthiness of results. As highlighted in Section 3.1.2, this research aims to be sensitive and adaptive to how peoples' information- and recommendation-seeking strategies may vary across tasks with different characteristics.

In cases where many potential information sources are identified within the user's social network, the approach presented here aims to help the user choose the most appropriate or trustworthy source of information given the characteristics of the information-seeking task. In doing so the aim is to be applicable and useful across a broader range of domains. This will be achieved by developing a detailed understanding of the source selection process in word-of-mouth recommendation, to be presented in Chapter 4.

3.2.3. Spam-resistant Information-seeking

A recent investigation (albeit journalistic, rather than scientific) (Walsh and Swinford, 2006) into 'review and rating spam' demonstrated how easily misleading reviews and ratings can be created on travel recommendation sites such as *TripAdvisor*⁶, by those with a vested interest in promoting a particular establishment. The investigation suggested that this form of manipulation is widespread; consequently recommender systems that base

⁶ <http://www.tripadvisor.com/>

recommendations on data that can be so easily falsified risk reducing the quality of their results (Josang, Ismail et al., 2007).

The use of social networks to support information-seeking makes the approach presented here less vulnerable to spamming, for the simple reason that each user is in the first instance only exposed to information or recommendations from people they know personally. This acts as a safeguard against manipulation of results, assuming that most users are unlikely to know others wishing to manipulate search indices on an ongoing basis, and at the expense of their acquaintances.

In the event that one individual persistently attempts to manipulate results, only those users who know the individual personally will be affected. These users will have the option of removing the individual from their social network (either virtually or in entirety!). The same benefits and safeguards do not apply to approaches based on social networks and trust propagation; by definition others beyond the immediate network will also be affected as trust relationships are propagated.

3.2.4. Openness to Additional Information

The approach presented here is oriented as much towards providing 'scaffolding' to support users in completing their information-seeking tasks, as it is toward providing solutions to their information needs. The aim is to augment rather than replace users' own assessments of members of their social networks as potential information sources. This is facilitated by the source-centricity of the approach and the use of social networks of known individuals.

For example, systems such as Hoonoh (Chapter 7) can make suggestions of potential sources of information on a particular topic; these can then be supplemented by users' own judgements of the suitability and trustworthiness of these individuals as information sources in the current task context. This can be beneficial in situations where additional knowledge about the appropriateness of sources is available to the user but not the system, or where the user wishes to be more selective about the choice of information source. It also provides a form of 'safety valve' in case of any discrepancies in how the system and the user perceive the trustworthiness of a source.

This contrasts with existing approaches such as collaborative filtering, where users do not have personal knowledge of the individuals upon whose preferences recommendations are based, and can only rely on reading reviews of suggested items (where available) as a means to assess the relevance of results.

In addition, members of his social network can provide the information seeker with knowledge to which they personally have access but that may never be available online (Kautz, Selman et al., 1997a). The approach developed in this research is limited in scope simply by the knowledge of members of one's social network, and the ability to infer the source most appropriate to the task. Whilst conventional approaches require items to be known to the system in order to be recommended to a user, this approach simply suggests (in the first instance) the most appropriate source from whom to seek further information, based on knowledge held about the characteristics of that source. Consequently, this approach allows for more graceful degradation compared to collaborative filtering, does not require maintenance of a central catalogue of items, and does not limit systems to operating in specific domains.

A further advantage of a social network-based approach is that additional follow-up questions about a topic can be addressed directly to known individuals as required. This may be conducted through any suitable medium (face-to-face, telephone, email) and does not need to be restricted to online communication through systems implemented as part of this research.

This ability conveys a number of benefits in addition to accessing information that is not available online: the information seeker may be able to fully communicate their needs and build understanding about their requirements in a way that might be hard to convey to a technical system; the information source may help to reformulate the problem where necessary (Cross, Rice and Parker, 2001); and the information source is likely to have a reasonable knowledge of the preferences of the information seeker (Sinha and Swearingen, 2001) and can tailor additional information accordingly.

A potential limitation of using only known social networks is a reduction in the number of information sources, compared to anonymous approaches. In order to overcome this issue, the option can remain to use unknown sources if necessary whilst accepting that these sources may provide fewer benefits.

3.2.5. Greater Reuse of Existing Data

In contrast to collaborative filtering systems, Hoonoh (see Chapters 5 and 7) can make use of existing data from a range of sources. Input data is still required in order to infer trust relationships with which to support information-seeking, but this data can take a broader range of forms and originate from many different sources. This aspect will be discussed in detail in Chapter 5. Less rigid data requirements may also make this

approach more adaptable to situations where available data may be too sparse for conventional statistical approaches.

Linden, Smith et al. (2003) outline limitations of traditional collaborative filtering that stem from its computational expense over large datasets, and demonstrate that computing taste overlaps is not scalable with very large numbers of users. The approach pursued in this research may be able to avoid this problem when it arises, by constraining information sources to those within the user's known social network rather than all other users of a system, thereby reducing the number of trust relationships that must be computed between users of the system. Whilst this may limit functionality in cases where users do wish to see results from unknown users, it will allow the system to scale more readily to large numbers of users.

Reuse of existing data to populate the system also reduces the amount of input required from users of the system and potential information sources. The approach presented here non-intrusively gathers information about the areas of knowledge of each member of a social network, from existing sources, allowing this information to be queried without requiring the active involvement of potential information sources or the sending of 'broadcast' messages to an entire group, which would likely increase information overload and quickly become aversive (Kautz, Selman et al., 1997b). This approach may also provide an incentive for individuals to make available data that can be used by the system, as information shared once can be reused many times by people they know.

3.3. Summary

This research proposes an approach to enhancing information-seeking on the Web through the use of social networks. Key characteristics of the approach are its source-

centricity, and its adaptivity to information-seeking tasks that have a range of characteristics. It has been argued that this approach brings a number of benefits: increased personal relevance, utility across a range of tasks, spam-resistant information-seeking, openness to additional information, and greater reuse of additional information.

However, to fully realise this approach it is necessary to better understand the dynamics of the information- and recommendation-seeking process among members of a social network, particularly how people assess the relevance and trustworthiness of information sources in tasks that extend beyond taste domains.

The following chapter reviews existing research in this area, identifies a number of limitations of this work, and presents an empirical study that addresses these limitations, thereby providing a richer understanding of the domain upon which the remainder of this research can be based.

4. Source Selection in Word-of-mouth Information-seeking

4.1. Background and Related Work

Word-of-mouth recommendation and referrals from others are powerful mechanisms for helping people acquire information and solve problems, in domains as diverse as finding piano teachers (Johnson Brown and Reingen, 1987) and successfully completing projects in the workplace (Cross, Parker, Prusak et al., 2001). Social networks of known individuals can serve as both a source of new information and as a filter to identify the information or items most relevant to one's specific needs (Granovetter, 1973),(Kautz, Selman et al., 1997a).

These processes have been extensively studied in a number of disciplines, particularly sociology, psychology, marketing and organisational sciences. In one of the earlier studies on the subject, Whyte (1954) provides an account of how interpersonal communication networks in local neighbourhoods can influence purchasing behaviour of domestic appliances. This study emphasised the existence of social networks that, through their role in information flow, can account for the non-random distribution of consumption patterns within the wider population. However, the work of Whyte (1954) was based on anecdotal evidence, and did not examine the nature of interpersonal relations between nodes in such networks or any effects these may have on the flow of information and subsequent purchasing decisions.

4.1.1. The Role of Weak Ties

When looking specifically at the relationship between the information seeker and an information source, one of the major themes in published work has been the notion of *strong* vs. *weak ties* in social networks, drawing on the work of Granovetter (1973). Whilst generally treated as discrete values of *strong*, *weak* or *absent*, *tie strength* is defined as a continuous variable stemming from a combination of amount of time, emotional intensity, intimacy and reciprocal services within a relationship. Importantly, it is posited that "the degree of overlap of two individuals' friendship networks varies directly with the strength of their tie to one another" (pp. 1360) (i.e. the stronger the tie between two individuals the greater the number of friends in common), and that a stronger tie correlates with greater similarity between two individuals.

Weak ties are considered more likely to act as 'bridges' between otherwise disconnected portions of the broader social network (supported empirically by Johnson Brown and Reingen, 1987). It is these weak ties that Granovetter found to play a key role in the diffusion of information to individuals who may not otherwise have been able to access it. Contrary to reasonable intuition, he found that weak rather than strong ties are more useful as sources of information about new jobs. This was attributed to the lower overlap between one's own social circle and those of others to whom one is weakly tied (i.e. a sufficient proportion of acquaintances were not shared). Consequently weak ties are more likely to be able to provide access to information about job opportunities that would be otherwise unavailable.

It is worth noting that Granovetter (1973) does not explicitly examine the way in which strong vs. weak ties affect the finding of a new job when elements of personal

recommendation and referral are involved; the study is simply concerned with access to information about job vacancies.

Johnson Brown and Reingen (1987) identify a shortage of empirical evidence to support the importance of weak ties in communication flows in social networks. They argue that existing studies are insufficiently general, tending to focus on the role of weak ties in just one setting. Furthermore, they cite later work by Granovetter (1983) that highlights how the 'strength of weak ties' argument has often been used more as a post-hoc rationalisation for empirical findings than as the focus of a systematic investigation.

4.1.2. The Role of Strong Ties

In addition to identifying shortcomings in the literature regarding the role of weak ties, Johnson Brown and Reingen (1987) also argue that there is potential for greater understanding of the role of strong ties in different aspects of word-of-mouth communication. The study they report seeks to provide empirical evidence for the 'strength of weak ties' argument of Granovetter (1973), whilst also examining the importance of strong ties in information-seeking and in influencing the decision-making of information recipients. Underpinning their work is a distinction between *relational form* and *relational content*. Relational form "refers to properties of the linkage between pairs of actors that exist independently of specific contents" (pp. 351); tie strength is one of these properties that make up relational form. Word-of-mouth recommendation information is given as an example of relational content.

Johnson Brown and Reingen make a subtle distinction in their work between the activation of ties for the flow of information in general, and active information-seeking through ties. The former can be thought of as 'did information flow through this tie?',

whilst the latter can be conceptualised as 'was this tie actively sought out as an information source?'

From a study of word-of-mouth information flow regarding piano teachers in a metropolitan setting, Johnson Brown and Reingen found that: strong ties and ties between homophilous individuals (i.e. those who have characteristics in common) are more likely than weak or heterophilous ties to be activated for the flow of referral information.

However, the hypothesis that "active information-seeking is more likely to occur from strong-tie than weak tie sources of referrals" (pp. 353) was not supported in the study. In fact, information was actively solicited from eighty six percent of weak ties used as sources, compared to active solicitation from only fifty percent of strong ties. This finding was attributed to the likelihood of incidental word-of-mouth communication increasing in line with communication frequency; therefore strong ties may be more likely to provide the required information in passing. It may be that where strong ties are unable to provide information in passing on a particular topic weak ties are actively sought instead.

Where referral information was provided by a strong tie it was perceived as more influential than referral information provided by weak ties. *Source credibility* is suggested as one explanation for the increase in perceived influence of information from strong ties, and a number of quotes are reported that suggest bases for this in factors such as *trusted opinions*, *valued recommendations*, and *knowledge of the field*. However, these factors are not investigated by Johnson Brown and Reingen, who do suggest that further

investigation of how attributes such as credibility influence the choice of information source may complement the findings of relational analyses such as theirs.

4.1.3. Influences on Choice of Tie-Strength

Duhan, Johnson, Wilcox et al. (1997) investigate how attributes of the information seeker (*prior knowledge*) and the task (*difficulty*, role of *instrumental and affective evaluative cues*) impact upon the use of strong or weak ties as information sources. Their study used a scenario-based approach but focused solely on the domain of medical services, specifically the search for recommended obstetricians.

Duhan et al. found that the greater the perceived difficulty of the task, the greater the chance that strong-tie sources would be sought for recommendations; this finding supported their hypothesis of a positive relationship between task difficulty and the seeking of recommendations from strong ties. Contrary to another hypothesis, it was found that a greater importance of affective evaluative cues in decision-making did not correlate with a greater likelihood of seeking strongly-tied recommendation sources. However, as hypothesised, a greater importance of instrumental evaluative cues in decision-making was found to correlate with a greater likelihood of seeking weak ties for recommendations.

Whilst the findings of Duhan et al. may appear to enhance our understanding of how task characteristics in particular impact upon the seeking of strong and weak ties as recommendation sources, their study has a number of limitations. The hypotheses investigated are based on a theoretical model formulated from previous research: however these hypotheses do not cover all possible relationships between factors present in the model, only certain relationships the authors predict to be of significance.

For example, the study predicts a relationship between task difficulty and recommendation-seeking from strong tie sources, but there is no comparable hypothesis testing a possible relationship between task difficulty and weak tie sources. In another example a positive relationship is predicted between the importance of instrumental cues and use of weak ties, without also examining possible relationships between instrumental cues and use of strong ties.

Consequently, it is not possible to conclude whether support for these latter two hypotheses was simply due to a greater chance of seeking recommendations *at all*, whether from weak or strong ties, as the design of the study is not sensitive to this. It is possible that other significant relationships exist that were not identified in the study but would invalidate the model. As a result, the study by Duhan et al. provides little evidence of the role of task characteristics in determining the use of strong or weak ties.

On this basis, it may be questioned whether relational form alone, and tie strength in particular, can provide an adequate, sufficiently granular, account of how people choose word-of-mouth information sources. In fact, attempting to explain source choice in terms of tie strength may represent a misapplication of the original research in this area. In Granovetter's (1973) work, tie strength is seen as a structural property that can influence information flow within networks, rather than a relational characteristic on which people base source selection decisions when actively seeking information. Consequently, tie strength may provide a rather blunt tool with which to understand source selection in information-seeking.

4.1.4. The Role of Source, Task and Individual Characteristics

A number of studies have moved beyond the broad strong/weak tie distinction and looked in more detail at the attributes of information sources that impact upon their selection by information seekers. Perhaps the largest body of work in this area concerns information-seeking within the workplace, from both human and non-human sources.

Workplace Studies

O'Reilly (1982) studied the frequency of use by welfare agency employees of a range of information sources, such as written documents, internal group members, and external sources. The impact of *source characteristics* (quality, accessibility), *task characteristics* (uncertainty, complexity) and *individual characteristics* (tenure, formal education, motivation) on frequency of use was investigated. In the context of this dissertation the most interesting findings relate to the source characteristics of quality and accessibility.

Accessibility of an information source was found to predict frequency of use for written documents (e.g. handbooks, procedures, memos, newsletters) and external sources but not human sources within the group. Further analysis found the frequency of use of group members to be a function of source quality, source accessibility, and the interaction between these factors. This interaction manifested itself in more frequent use of high quality, low accessibility sources than low quality, high accessibility sources, with sources of high quality and high accessibility being preferred.

O'Reilly acknowledges that *quality* is a subjective concept. He uses attributes such as *relevance*, *specificity*, *accuracy*, *reliability* and *timeliness* to define a more general notion of *information quality*, and it is at this higher level that the analysis is conducted.

Characteristics such as the expertise of the source are not explicitly included, although aspects of this factor may be somewhat accounted for by relevance and accuracy.

Borgatti and Cross (2003) investigated one team of scientists and another of researchers to specifically examine the impact of different factors on their choice of human information sources. Whilst O'Reilly's study is broad in terms of factors analysed (source, task and individual characteristics), Borgatti and Cross present a model that encompasses more stages of the information-seeking process.

They hypothesise that "the probability of seeking information from another person is a function of (1) knowing what that person knows; (2) valuing what that person knows; (3) being able to gain timely access to that person's thinking; and (4) perceiving that seeking information from that person would not be too costly." (pp. 432). The notion of source quality is also present in the model of Borgatti and Cross but under the label of 'valuing what a person knows'.

Results of the study support the hypotheses that *knowing*, *valuing* and *access* predict the use of a source for information-seeking; *cost* was not found to be significantly related. In addition, Borgatti and Cross found that *knowing* and *access* mediated the effect of *source proximity* on information-seeking, supporting their assertion that the effect of proximity in intentional information-seeking is indirect. This suggests that people ask others who are proximal not specifically because they are proximal, but rather because by virtue of being proximal they are easily accessed and the information seeker is more aware of what knowledge they may have.

Cross and Borgatti (2004) report a similar study that examined the impact of the source-seeker relationship on information-seeking. Through interviews with managers in a

business consulting practice they identified four characteristics that were hypothesised to predict information-seeking: *awareness* of a potential source's expertise, *timely access* to the source, the *safety* of the relationship and willingness of the source to *cognitively engage* with the problem. A model based on these characteristics was then formulated and tested.

It was found that *awareness*, *timely access* and *engagement* were all predictors of source choice in information-seeking, however the same was not true for *safety*. These findings highlight that simply knowing who has knowledge or expertise in a topic is not sufficient in selecting an information source, as one must also be able to access a source who must also be willing to engage in problem solving. This study also provides some support for the findings of Borgatti and Cross (2003), as the *knowing/awareness* and *access* factors were found to be significant in both studies.

Morrison and Vancouver (2000) found that, in a sample of career-early aerospace engineers, *expertise* and *accessibility* of information sources both predict the likelihood of that source being used. Of these two factors, expertise was found to have the greatest impact. It is worth noting that the participants in Morrison and Vancouver's study were asked to rate information sources from a fixed list (supervisor, friend, colleague, mentor, documents) rather than sources they identified themselves. Despite this limitation, the results strongly support the findings of Borgatti and Cross (2003) and Cross and Borgatti (2004) relating to accessibility of information sources. The outcome related to perceived expertise of the source supports Borgatti and Cross's (2003) finding that perceived value of a source predicts use of that source.

McDonald and Ackerman's (1998) study first introduced in Chapter 2 distinguishes between *expertise identification* ("the problem of knowing what information or special skills other individuals have") (pp. 317) and *expertise selection* ("appropriately choosing among people with the required expertise") (pp. 317), also in a workplace information-seeking context. Expertise identification appears to be closely related to the *knowing* identified in many of the studies reported above.

McDonald and Ackerman identified a highly specialised role in their field study, that of the *expertise concierge*, who maintains a sophisticated mental model of members of the organisation and what they know. This can then be used to refer information seekers to potential sources, easing the task of expertise identification.

Expertise selection was found to be influenced by three mechanisms, which bear some relation to those found in the studies discussed above: *organisational criteria*, the *load on the source*, and *performance*. The first of these, as the label suggests, refers to organisational aspects outside the scope of this research. Secondly, sources with lower day-to-day workloads, or who had not been heavily used as sources, were more likely to be approached for help. This would appear to be related to the accessibility construct identified in studies discussed above. Lastly, *performance* is broken down into sub-components of *problem comprehension*, *providing a suitable explanation*, and *attitude*, which together appear to form a construct very similar to Cross and Borgatti's (2004) *willingness to cognitively engage* factor.

Interestingly, in this study expertise is treated as a general notion that people seek out in order to perform their work functions, rather than one component factor in predicting information-seeking, as in studies such as Morrison and Vancouver (2000).

Non-workplace Studies

Whilst the studies discussed so far have provided a range of explanations for relational and attributional factors that influence source choice in information-seeking, their focus is limited to workplace settings. Fewer studies have been conducted in less formal domains where the requirements and priorities of the situation may differ, resulting in an emphasis on different factors in selecting information sources.

For example, in taste domains people may be more oriented towards choosing sources who share similar tastes, perhaps in favour of domain experts. In this case our social networks of known individuals may prove particularly helpful. Due to homophily (McPherson, Smith-Lovin et al., 2001) we may (amongst other dimensions) be expected to also share with them many tastes. Despite these factors, the literature on information-seeking and source selection in taste domains is relatively limited.

Bonhard and Sasse (2005) recognise the need for greater research in this area, and report on a qualitative study of recommendation-seeking in taste domains. Participants in the study were asked questions about how they chose services such as plumbers, lawyers and doctors, which may be seen as less taste-oriented, however the authors do strongly emphasise taste domains as the focus of the study, and consequently report many results related to domains such as music, books and films.

A major dimension identified in Bonhard and Sasse's study, and the primary dimension in their resulting model, is *objective domains* vs. *taste domains*. Objective domains are defined those in which "items are characterised by measurable and comparable specifications" (pp. 260), with examples given such as electronic goods, computer

hardware and software, and cars. Examples given of taste domain items are music, books, films and restaurants.

Advisor expertise is the only relational or attributional factor present in the model that is seen to affect the weight given to advice in objective domains; the other factors influencing advice seeking in objective domains both relate to characteristics of the information seeker. The part of the model addressing taste domains contains many more factors addressing aspects of the task, the relationship and the information source. These factors are: *risk*; *past experience* with the source; *source reputation*; *advisor expertise* (present in both objective and taste domains); and whether the source is or is not *known personally* (written sources such as reviews in magazines were included in the study). When it comes to making a decision about a piece of advice that has been received, *trust* and *reliance* are seen to have an impact.

Risk in Bonhard and Sasse's model is seen to include financial risk, and whether the domain of recommendations is oriented toward *experiences* (e.g. cinema trips or restaurants) or *consumption* (e.g. books or CDs). A greater financial risk was generally found to be associated with more thorough research before making a decision, although it is not apparent from this study which factors are most taken into account in higher financial risk situations. The study found that choosing *experienced goods* was seen as higher risk than *consumed goods* (the authors concluded this was due to being able to return items such as books and CDs), and consequently in such situations people chose advisors seen as trustworthy, "namely known recommenders and or those with a good track record" (pp. 261).

In the domain of *services*, such as hairdressers, plumbers, lawyers and doctors (it is not apparent whether these are considered objective or taste domains), Bonhard and Sasse found that information seekers sought recommendations from friends irrespective of whether those friends had expertise in the domain in question. They conclude that in such cases people do not aim to get reliable information about the quality of a particular service, but simply to get reassurance about whether a service is OK.

Past experience, *source reputation* and *advisor expertise* are grouped together under *trust* and *reliance*. Past experience refers to the likelihood that good recommendations from a source in the past will lead to reuse of the source in the future. Reputation and expertise are not explicitly defined to any greater extent, except to say that they can both increase trust in a first time encounter. The factors that determine a source's reputation or expertise are not specified, nor is the relationship between past experience and reputation.

In Bonhard and Sasse's model, the importance of knowing a source personally is attributed to two factors: *taste overlap*, whereby the recommendation seeker knows the sources has similar tastes, and *mutual knowledge*, which allows a source to provide recommendations even where they do not share similar tastes.

Whilst the work of Bonhard and Sasse provides some useful insights into choice of source in taste domains, the model has a number of limitations. Firstly, the study on which the model is based is oriented towards taste domains and does not systematically investigate objective domains, even though this distinction appears to have emerged from the data and is present in the model. Perhaps as a consequence, the model is relatively underspecified regarding choice of source in objective domains compared to the degree

of detail in taste domains. There does not appear to be a clear conception of objective domains in the study, as some examples given (e.g. cars) could equally be considered as taste domain items.

Secondly, many factors are identified in the study and represented in the model, however some of these (e.g. reputation, expertise) remain poorly defined and their relationship to other factors unclear (e.g. past experience, reputation). Furthermore, it is not apparent to what extent the model is predictive of which members of a network would be chosen as information sources in a specific scenario, and the authors do not provide evidence to demonstrate such predictive validity.

These shortcomings of Bonhard and Sasse's model may reflect the inherent complexity of the domain, or simply that the model would benefit from clarification in some areas. Either way, in its current form the model is too complex and insufficiently specified to enable it to be operationalised in technical systems.

4.1.5. Summary

The literature reviewed above provides indications of how the source selection process in information-seeking operates. A number of recurrent themes are present across the reviewed studies, such as the accessibility of sources and their perceived quality.

The diagram shown in Figure 2 below provides a representation of the information-seeking process from a source identification and source selection perspective, showing factors identified in studies reviewed in this chapter as having an impact on source selection in word-of-mouth information-seeking.

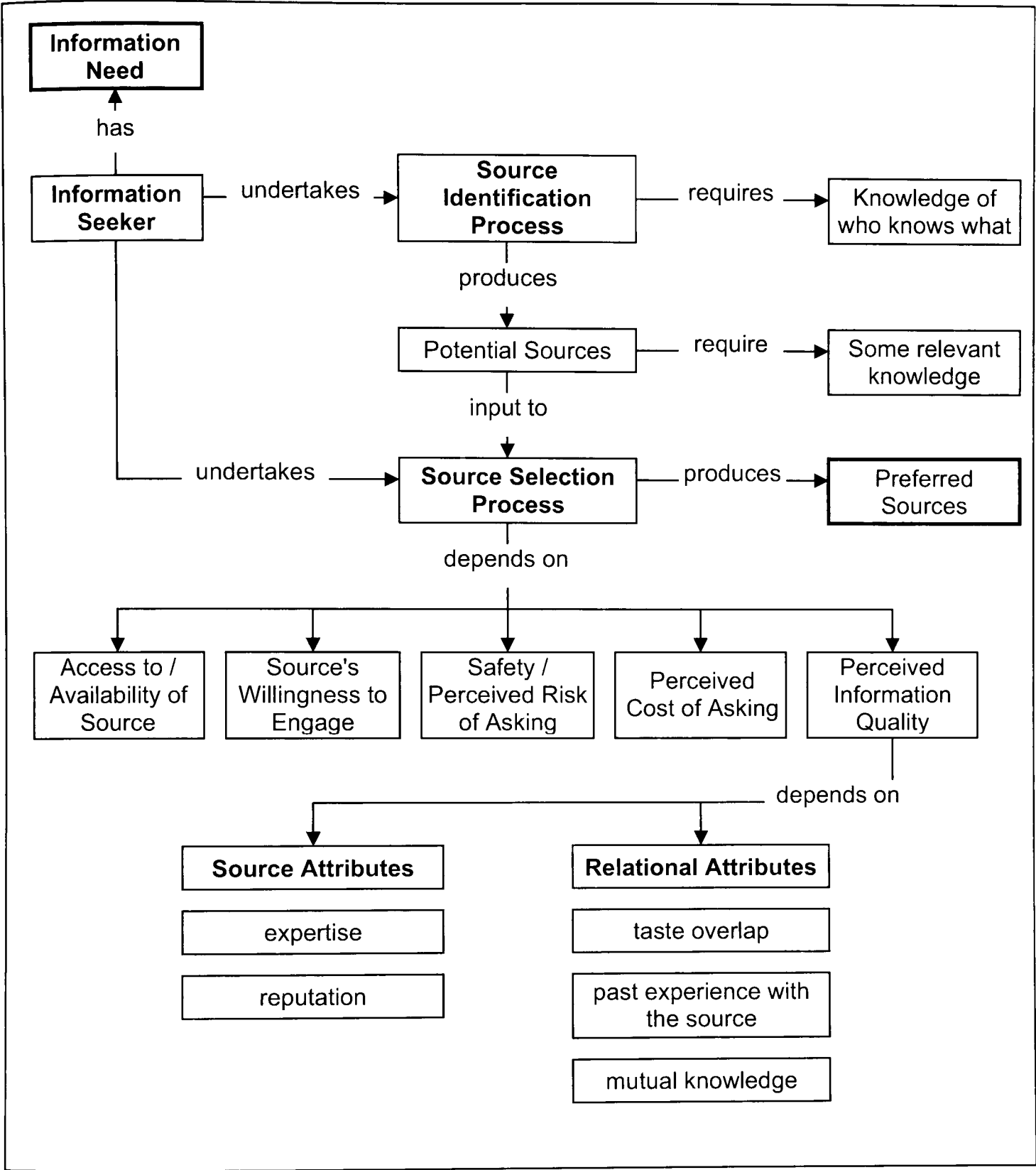


Figure 2. The information-seeking process from a source identification and source selection perspective.

The studies in which the source and relational attributes shown were identified are shown in Table 1 below. Quality is included in this table to aid comparison between O'Reilly's work and that of others; however, Figure 2 above reflects the notion of quality as a higher-level construct that subsumes more specific factors.

	O'Reilly (1982)	Morrison and Vancouver (2000)	Bonhard and Sasse (2005)
Source Attributes			
"quality"	x		
expertise		x	x
reputation			x
Relational Attributes			
taste overlap			x
past experience with the source			x
mutual knowledge			x

Table 1. Source and relational factors identified in existing literature as affecting perceived information quality

The literature on source selection in information-seeking is dominated by studies from workplace settings that deal primarily with information-seeking in job-related tasks. Studies investigating source selection in less informal and more taste oriented domains are less widespread. Whilst Bonhard and Sasse's (2005) model does distinguish objective domains from taste domains, this factor is not systematically varied in the study on which the model is based and the findings remain oriented towards source selection in taste domains.

Overall, a picture of the source selection process does not emerge that is sufficiently consistent or generalisable to serve as a basis for implementing technical systems that support the selection of information sources within one's social network.

In order to establish some general principles from which the source selection process may be modelled, a further investigation is required that enhances our understanding of how people select information sources across a broader range of tasks, in domains not only mediated by taste. To address this need an empirical study was carried out to explore: from whom people seek information and recommendations in different

scenarios; the factors that underlie their decisions about the trustworthiness of this information; and how the influence of these factors varies across different types of task.

4.2. Study of Source Selection in Word-of-mouth Information-seeking

This study addresses research Questions 1-3, introduced in Chapter 1:

1. How do people choose information and recommendation sources from among members of their social network?
2. Which factors influence judgements about the relevance and trustworthiness of these information and recommendation sources?
3. How do the characteristics of the task being performed affect these judgements?

Previous work in the area, as discussed above, does not provide a sufficiently comprehensive and consistent account of the information-seeking and recommendation-seeking process from which hypotheses can be derived and tested using quantitative methods. Therefore by necessity this study is exploratory in nature and qualitative in methodology. The aim is to identify central themes and factors in the decision-making process and gain insight into how the influence of these factors varies across different types of tasks, in order to identify general trends that can be operationalised in technical systems.

4.3. Design

The study consisted of semi-structured interviews in which participants were presented with a series of fictional recommendation-seeking scenarios and asked a number of open-

ended questions exploring their decision-making process when selecting an information source.

4.3.1. Pilot

A pilot was conducted with three participants (who were not included in the main sample) to test the experimental protocol. This led to refinement of the interview script in order to ensure the results produced by the study would be sufficiently relevant to the research questions. In particular the open-ended questions used in the study were modified in order to be more structured, as the pilot had demonstrated that participants did not always understand how to respond to very open-ended questions.

4.4. *Method*

4.4.1. Participants

Twelve participants were recruited to the study using opportunistic sampling. Participation was voluntary, and no payment was received for taking part in the study. All participants were staff or students at The Open University, and varied in age from mid-20s to mid-50s. Seven participants were male and five were female. Whilst the majority of participants were British, participants from Germany, the Netherlands, New Zealand, Ukraine, and the USA were also present in the sample.

4.4.2. Procedure

The study consisted of one semi-structured interview with each participant, on a one-to-one basis, in person. Interviews lasted between 16 and 60 minutes, varying according to the participant's engagement with the topic. After being given general instructions about

how the interview would proceed, the participant was read in turn each of four hypothetical information- and recommendation-seeking scenarios (reproduced in Table 2 below) and asked to imagine themselves in this situation.

The scenarios used in the study were constructed by the researcher, and designed to closely represent everyday tasks and situations in which recommendations might be sought from members of one's social network. This contrasts with studies by authors such as O'Reilly (1982) where similar issues are investigated, but specifically in a workplace setting. It is not apparent how applicable such findings are outside that particular domain. The scenario-based approach bears some similarities to that used by Duhan, Johnson et al. (1997); however, in this case each participant was presented with multiple scenarios covering a range of domains, compared to Duhan et al's use of one scenario in a single domain.

Number	Domain	Text	Criticality	Modality
1	Plumber	"You move into a new house that requires renovation, including some substantial plumbing work. Who would you ask about recommended plumbers?"	High	Locating
2	Back Pain Treatments	"You are suffering from moderate and ongoing back pain and need to find some ways of getting it treated. Who would you ask about recommended ways of getting it treated?"	High	Exploring
3	Business Hotel	"You are travelling to Madrid on business and need to find a hotel to stay in during your visit. Who would you ask about recommended hotels?"	Low	Locating
4	Holiday Activities	"You are planning a holiday to the east coast of the USA and need to find some information about how to spend your time there. Who would you ask about recommended activities?"	Low	Exploring

Table 2. Recommendation-seeking scenarios used in interviews with participants

The tasks described in the scenarios were varied along two dimensions: *task modality* and *task criticality*. Making up the task modality dimension, two of the scenarios (*plumber*, *business hotel*) described *locating* tasks, whilst two (*back pain*, *holiday activities*) described *exploring* tasks, as defined in Heath, Dzbor and Motta (2005). Locating tasks are those where the user is seeking a specific item or piece of information that is believed to exist, and the challenge is to identify an appropriate option or solution from among many. In contrast, exploring tasks are those where the user is attempting to develop a broad picture or understanding of a domain; the challenge in this case is to gather a representative range of perspectives from which later decisions may be taken.

Task criticality was defined as the degree of risk associated with a poorly chosen item or solution. This dimension was represented by two scenarios where the task was seen as low-criticality to the information seeker (*business hotel* and *holiday activities*), and two where the task was seen as highly critical (*plumber* and *back pain*).

The study was mindful of possible effects of domain (e.g. tourism, healthcare) and locality of task (for example, tasks based on information about the local area vs. information about distance locations), but these were not systematically varied in the study.

After being read each scenario, the participant was asked a series of questions, which can be paraphrased as:

- *From whom they would seek a recommendation?*
- *Was there anyone they would not ask?*
- *What were the reasons for these decisions?*

These questions made up a common script used by the experimenter (see Appendix A⁷), which provided a general structure for the interviews. This structure was broadly followed, however in line with the exploratory nature of the study deviation by participants was permitted in order to capture as rich an account of the decision making process as possible. Participants often provided lengthy responses which rendered later questions irrelevant, in which cases these questions were skipped by the experimenter. Asking participants if there was anyone they would not ask provided an opportunity for participants to elaborate on their source selection rationale, and often provided a richer picture of their decision-making process.

⁷ The script included questions about the effect of poor recommendations on future recommendation seeking from that source. Responses to these were not sufficiently in scope for this research and consequently were excluded from the analysis.

It was emphasised to each participant that there were no right or wrong answers to the questions asked by the interviewer, but simply that the research was interested in how they approach the problems presented in the interview.

Participants were not limited to specifying information sources within a certain proximity in their social network. Some did ask for clarification regarding whether they could cite sources not known to them personally, and some actively cited other sources such as the Web, however these cases were rare. Participants were also not constrained to citing sources with any particular tie-strength, as this was not a variable in the study. This allowed for examination of the salient properties of the information source or the interpersonal relationship as these impacted on the task in the scenario, without this being obscured by questions of tie-strength.

Participants were also asked to describe any analogous recommendation-seeking scenarios from their own experiences which came to mind in the course of the interview, and describe to their decision-making process on these occasions. Data from these accounts was included in the analysis.

Audio recordings of the interviews were made and transcribed to form the basis for the analysis.

4.5. Analysis

Following the methodology described in Smith (1995), inductive analysis of the transcripts was carried out to identify themes in respondents' decision-making.

Each transcript was systematically analysed to identify factors that determined from whom respondents would seek recommendations. The factors identified across all

transcripts were aggregated into a master list, from where they were grouped into a list of initial themes which was grouped again to produce the super-ordinate themes described below. The master list and initial themes are reproduced in Appendix B.

4.6. Results and Discussion

Five factors were identified that influenced participants' choice of sources for word-of-mouth recommendations, and the trust and confidence they had in information from these sources. Definitions of these factors are provided below, followed by frequency data and illustrative quotes taken from transcripts of the interviews. From now onwards these factors will be referred to as 'trust factors'. Factors related to practical aspects and diversity of responses were also raised, and are included in the Appendix; however, these were not included in the analysis as they do not relate to trust and relevance issues.

4.6.1. Trust Factors: Definitions

- **Expertise:** the source has relevant expertise of the domain of the recommendation-seeking; this may be formally validated through qualifications or acquired over time.
- **Experience:** the source has experience of solving similar scenarios in this domain, but without extensive expertise.
- **Impartiality:** the source does not have vested interests in a particular resolution to the scenario.

- **Affinity:** the source has characteristics in common with the recommendation seeker, such as shared tastes, standards, values, viewpoints, interests, or expectations.
- **Track Record:** the source has previously provided successful recommendations to the recommendation seeker.

Note that *expertise*, *experience* and *impartiality* relate to relationships between an information source and the topic of the recommendation-seeking (these are person → topic factors), whereas *affinity* and *track record* capture a relationship between the source and recommendation seeker (these are person → person factors).

4.6.2. Trust Factors: Illustrative Quotes

The following quotes from participants in the study illustrate the five trust factors:

Expertise

"I would probably go and ask my friend who is a plumber or my friend who is a gas fitter, working on the principle that their domain expertise, their knowledge, is in a similar area."

Quote 1. Participant ID 16, Plumber scenario

"Maybe I would immediately approach my doctor in the surgery where I'm registered, and ask his advice. ...I wouldn't be confident that the advice is reliable...from the people who I don't know as specialists in the area."

Quote 2. Participant ID 10, Back Pain scenario

Experience

"I guess it depends on the location of the flat where I lived. If it was somewhere near to my parents I'd probably ask them first, for their advice, because they've got more experience, they've met people in the past who've done good jobs for them etc. etc."

Quote 3. Participant ID 05, Plumber scenario

"People I know in the area, it's good to have word-of-mouth, you know they've got experience good or bad."

Quote 4. Participant ID 14, Plumber scenario

Impartiality

"...with travel agents you'd have to question what they were promoting to you - is it because they get commission?"

Quote 5. Participant ID 08, Holiday Activities scenario

"Who wouldn't I ask? [I have] no specific examples. Actually its travel agents, as they're trying to sell you something; people who have no personal relationship to me and are interested in selling a product."

Quote 6. Participant ID 16, Holiday Activities scenario

Affinity

"There is someone I would not ask [for] recommendations, who it would probably help to speak with... they have been to the States this summer and previous times... but ... because we're different persons she cares about different details than me... and adding to is that I don't think we have the same style in things we are after, so I wouldn't be urged to ask her advice."

Quote 7. Participant ID 17, Holiday Activities scenario

"[I] may not ask people who I don't feel comfortable with, who haven't got the same values as me, or have a completely different lifestyle that I don't relate to."

Quote 8. Participant ID 12, Plumber scenario

Track Record

"I looked on the internet yesterday about going to see a masseur, but they were too expensive so I'll go back to [ask] my sister as I had a good experience with [recommendations from] her before."

Quote 9. Participant ID 07, Back Pain scenario

"Like the plumbing one [I wouldn't ask] someone who'd given me bad recommendations of hotels in the past."

Quote 10. Participant ID 16, Hotel scenario

4.6.3. Trust Factors: Occurrence Frequencies

Whilst the goal of the analysis was not to produce quantitative results for statistical analysis, it is useful to examine the frequencies of occurrence of the different trust factors in participants' explanations for choosing a particular recommendation source. As shown in Figure 3, expertise, experience, and affinity occurred most frequently, with relatively low occurrences of the impartiality and track record factors.

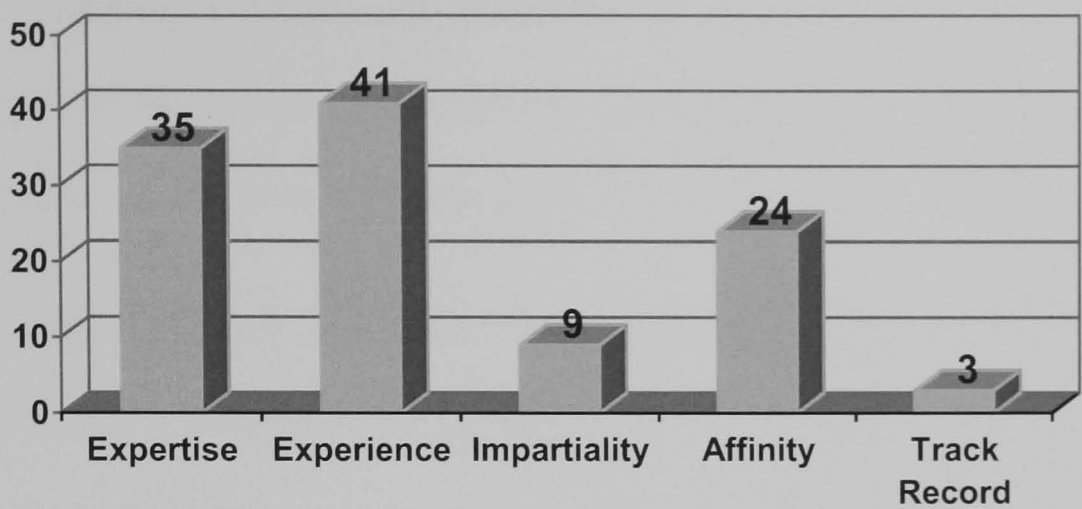


Figure 3. Response frequencies for each factor, summed across 4 scenarios in each of 12 interviews⁸

4.6.4. Trust Topicality

It is worth noting that whilst the factors expertise, experience, and impartiality were clearly domain specific and therefore topical in nature, the study did not give a strong indication of affinity as a topical factor, but rather as a more general construct. This may seem counter-intuitive at first as this aspect of affinity contrasts with taste, which is generally treated as a domain-specific characteristic. The relationship between affinity and taste is explored in the following section, along with a general discussion of how the findings relate to previous work in the area.

4.6.5. Relation of Trust Factors to Previous Work

Comparing the results of this study to the findings of previous research it is apparent that whilst some commonalities exist some novel trust factors have been identified. *Expertise* was identified as a factor in source selection by Morrison and Vancouver (2000) and

⁸ Therefore giving a maximum frequency of 48

Bonhard and Sasse (2005). Bonhard and Sasse also identified how *past experience* with a source can affect future use of that source for information-seeking, as can taste overlap between the information source and information seeker.

The affinity factor identified in this study appears to be relatively novel. One reason for this not having been previously identified may be that outside the formal roles and structures of the workplace there may be greater potential for exercising personal discretion in selection of sources, increasing the use of affinity relative to other trust factors. Furthermore, the tendency for existing studies to examine either taste domains or workplace expert finding may explain why the more universal notion of affinity has not been previously recognised.

It appears that affinity may be crucial where subjective recommendations are sought rather than simply factual information, a conclusion consistent with the findings of Bonhard and Sasse (2005) regarding taste domains. However, the data obtained in the study (e.g. Quotes 7 and 8 above) indicated that affinity represents more than simply shared tastes and is in fact domain-independent. In addition to style and taste, affinity appears to encompass more universal traits such as similar outlooks on life, values, lifestyle, expectations and attention to detail. Whilst affinity and taste are no doubt related in some way, the results of this study suggest that they are not interchangeable. In fact, shared tastes would appear to be one sub-component of the broader notion of affinity, which can be thought of as 'taste++'.

The study reported here did not identify a specific role for mutual knowledge or reputation, both of which were identified by Bonhard and Sasse. However, Bonhard and Sasse do not adequately define the concept of reputation, which may simply reflect a

personal or social perception of the quality of information from a particular source. Reputation may in fact represent an aggregate measure of factors identified in this study, particularly expertise, experience, impartiality and track record. Any role of affinity would likely depend on whether a personal or group level definition of reputation was adopted.

In contrast to previous research, the study reported here identified relevant experience as a key factor in determining the trustworthiness of or confidence in an information source, along with the source's impartiality with respect to the domain of the task.

4.6.6. General Trends in Application of Trust Factors

Individuals did vary in the source selection strategies they reported, however some general trends emerged, most significantly that the emphasis given to each of the trust factors varied according to the characteristics of the recommendation-seeking task. These trends are examined in the following sections.

4.6.7. Effects of Criticality, Subjectivity

In tasks perceived as highly *critical* (e.g. the back pain scenario), emphasis was placed on externally validated '*expertise*', as illustrated by Quote 1 and Quote 2 above. This finding is consistent with the claims of Dieberger, Dourish et al. (2000) that "some domains depend more heavily on expert recommendations" (pp. 43).

In less critical tasks respondents were less selective. Some participants indicated a particular willingness to seek information from a broad range of sources in less critical situations, on the basis that information from less trusted sources could be filtered or disregarded later if necessary, as illustrated by the quote below.

"My view is I gather everything from everybody and filter it, so I wouldn't be averse to asking people who maybe wouldn't like the same holiday, I'd still be prepared to take on board what they recommended, because I'd then filter it out, rather than not taking it."

Quote 11. Participant ID 12, Holiday Activities scenario

Where tasks were perceived to have an objectively correct solution, respondents also widely cited '*expertise*' or '*experience*' of the recommender as influencing their choice. However, where suitable solutions were more *subjective* (such as in the holiday activities scenario), respondents emphasised the '*affinity*' factor. Some participants indicated that they would reject sources with highly relevant experience if there was not an affinity between themselves and that source, as illustrated by Quote 7 above.

These results suggest that the criticality of the task and the subjectivity of possible solutions were of primary importance in determining which trust factors were emphasised. In scenarios seen by participants as more critical, greater emphasis was placed on the recommendation source having relevant expertise. In contrast, in scenarios in which potential solutions were seen as more subjective, participants placed greater evidence on sources with which they shared a strong affinity.

4.6.8. Effects of Task Modality

Effects of task modality (i.e. locating vs. exploring) were not readily apparent in the data. This may indicate that sources are chosen in the same way irrespective of modality. However, it is also possible that variation in criticality of the tasks and subjectivity of solutions masked any such effects in this study.

4.6.9. Domain of Task and Nature of Relationship

Respondents indicated that they would choose information sources with '*expertise*' or '*experience*' appropriate to the domain of the task (e.g. a doctor in the back pain scenario). However, any variation in how the trust factors are employed across domains such as tourism and healthcare is attributable to factors such as the criticality and subjectivity of the task, not to differences in strategy that are specific to particular domains.

Close family and friends were often cited as sources. Whilst trust factors such as '*affinity*' and '*track record*' likely contribute to this finding, it is also probable that respondents cited these sources for practical reasons; they are easily accessible, and the seeker can better assess their suitability to give recommendations in a particular domain. The precise nature of the relationship between respondent and the source they chose did not appear of great importance. Practical factors such as the source being a gatekeeper to others (as a family doctor may be), and the social acceptability of asking someone were also mentioned.

4.7. Conclusions

This chapter reports on an empirical study examining how people seek recommendations from members of their social networks, across a range of scenarios. The study demonstrates that people make detailed and complex decisions when identifying sources of recommendations, and assessing the trustworthiness of such sources. Furthermore, these decisions take into account a detailed knowledge of potential recommendation sources.

Analysis of the data identified five factors that influenced from whom participants would seek recommendations, and how trustworthy these sources would be perceived to be: *expertise, experience, impartiality, affinity, track record*.

The specific factors on which source selection decisions were based varied according to the characteristics of the task. In particular the criticality and subjectivity of the task were found to influence the factors most attended to in a given scenario.

Whilst providing support for a number of findings from existing research, the results of this study make a number of novel contributions: they provide results that may generalise more readily, as a range of scenarios were used beyond purely workplace or taste domains, and these were supplemented by participants own accounts; they expand upon previous research by identifying new factors that influence source selection, thereby further unpacking the notion of source quality.

These findings address Research Questions 1-3, by identifying how people choose information and recommendation sources from among their social network, the factors that influence judgements of the relevance and trustworthiness of these sources, and how source selection decisions vary according to the characteristics of the task. The next chapter outlines how these findings will be utilised to develop technical systems that support social network-enhanced information-seeking.

5. Technical Approach and Architecture

The goal of this research is to enhance information-seeking on the Web using social networks as trusted information sources. Chapter 3 outlined at a conceptual level the approach being taken in this research, whilst Chapter 4 provided an empirical basis for subsequent components of the research.

This chapter gives an overview of the technical approach adopted in the research, which can be briefly summarised as: operationalising the findings of the study in Chapter 4 as computational algorithms; collecting data from distributed sources as input to these algorithms; using the algorithms to generate metrics that represent trust relationships pertinent to word-of-mouth information- and recommendation-seeking; and using these trust metrics as input to a Web-based system that supports source-centric information-seeking.

5.1. Architectural Overview

Figure 4 below provides a high-level overview of the technical architecture developed in the course of this research.

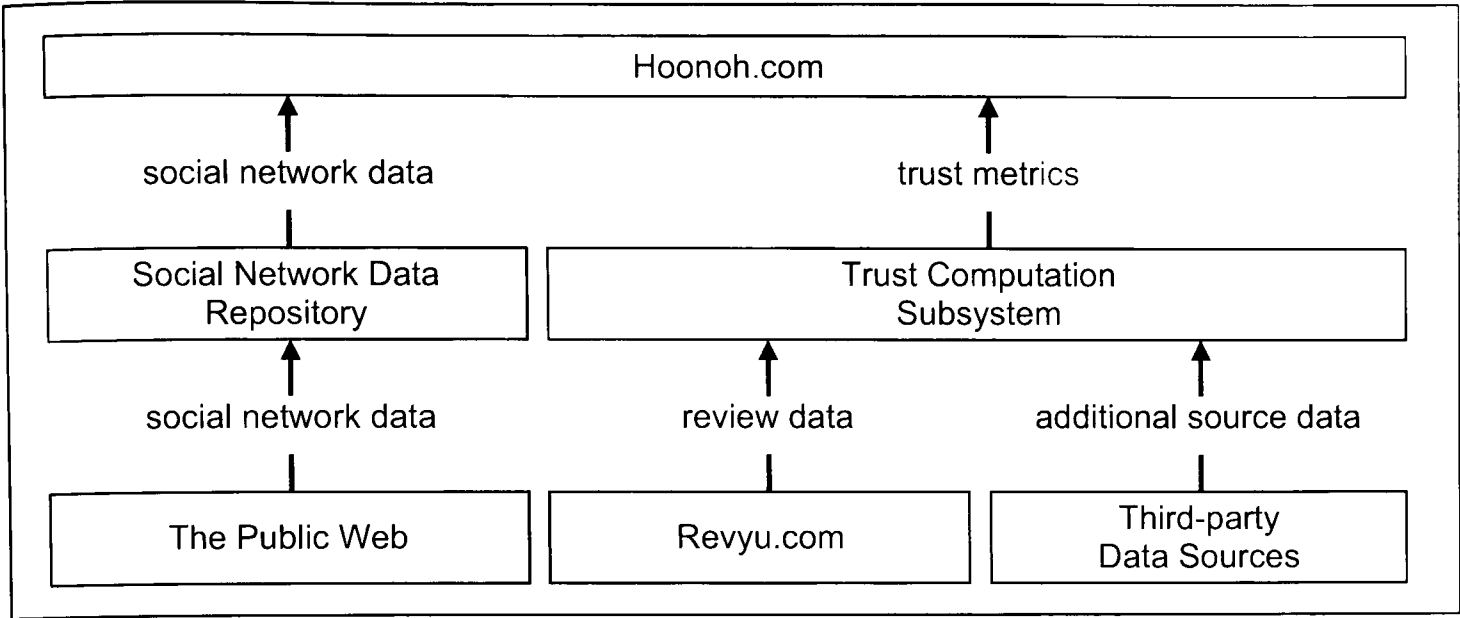


Figure 4. Overview of the technical architecture

A distributed technical approach has been adopted whereby data is acquired from a range of different sources as the basis for generation of trust metrics, and loosely coupled systems are provided that exploit this data. Separation of concerns in this fashion emphasises a distributed, Web-oriented approach that enables users to benefit from social network-enhanced information-seeking based on data they have already provided to other systems on the Web.

At an architectural level, the goals of the research could have been addressed using a single unified system. However, such an approach would have a number of disadvantages. All potential end users would be required to provide substantial amounts of data to the system, when this may already be available online in other locations. This could lead to unnecessary duplication of data. Furthermore, requiring all users to adopt a single system for both data input and social network-based information-seeking creates potential barriers to uptake, as users may be reluctant to adopt a new system that requires a high initial investment in order to be useful, reducing the likelihood of such a system reaching critical mass.

Existing and well-established systems such as *Del.icio.us*⁹ have some characteristics that are relevant to this discussion. *Del.icio.us* allows users to create a 'network', which consists of a list of other users. All bookmarks from members of this network are then aggregated for the user, providing an overview of current activity and topics of interest. Whilst pertinent to the discussion, this functionality does not address the aims of this research, as no trust element or task specificity is used to determine ranking of relevant items. Furthermore, *Del.icio.us* is item-centric rather than source-centric, and operates solely on Web pages rather than a range of online and offline items.

5.2. Computing Trust Relationships

This research has adopted an automated approach to computing trust metrics using existing data sources wherever possible, in order to minimise the effort required by users to bootstrap the system. An alternative approach would be to ask users to rate their own trustworthiness as information sources across a range of domains or assess members of their social network on the same basis. This approach was rejected for the following reasons:

- Manual provision of trust ratings was deemed unnecessarily onerous for users;
- Manual provision of trust ratings would require a comprehensive yet manageable list of topics or domains against which each person would be rated; by definition this scales poorly to the full range of topics on which users might wish to seek information or recommendations;

⁹ <http://del.icio.us/>

- Research has shown that individuals have a tendency to discount others opinions relative to their own (Yaniv and Kleinberger, 2000), so the accuracy of this method may be questioned.

5.2.1. Choice of Trust Factors

The empirical study reported in Chapter 4 found that 'experience', 'expertise' and 'affinity' were the most frequently cited factors influencing source selection in information-seeking. Priority in this research has been given to generating trust metrics based on these three factors, for the following reasons

- Having been cited most frequently in the empirical study it is reasonable to conclude that they collectively account for most variation in source selection decisions;
- The benefits provided by integrating 'impartiality' and 'track record' metrics may not justify the costs of computing these metrics;
- How a user's impartiality with respect to different domains may be computed, and based on what data, is unclear;
- Whilst track record did emerge as a theme in the empirical study, it may be hypothesised that a poor track record will over time result in lower affinity between two individuals, therefore separate computation of track record metrics is deemed unnecessary.

The sections that follow describe the data requirements for the algorithms that generate these metrics, and identify sources from which such data is obtained. Chapter 7 presents the algorithms themselves.

5.2.2. Data Acquisition from Distributed Sources

In order to compute experience, expertise and affinity trust metrics two basic types of data are required: data that connects people to domains or topics, from which experience and expertise metrics can be computed; and data that connects people to other people, from which affinity metrics can be computed.

The APIs of many so-called 'Web2.0' O'Reilly (2005) services such as *Amazon*¹⁰, *Del.icio.us*, *Flickr*¹¹ and *Facebook* provide data that may address some of these requirements. For example, keyword tags that people have used to annotate photos or bookmarks may indicate domains in which they have experience, whilst reviews of items on *Amazon* may provide a basis for computing affinity scores between users using collaborative filtering-style approaches.

Some use is made of data from these services, such as tagging data from the social bookmarking site *Del.icio.us*, as will be detailed in Chapter 7. Because tagging is unconstrained in the terms that can be used, tagging data has the potential to provide evidence of an individual's experience across an infinite number of domains, which would not be possible if a fixed topic list and manual ratings were used.

However, the data available through services such as *Del.icio.us* and *Amazon* is limited in a number of ways that affects its utility in this research, particularly in the extent to which reviews, tags and social network data can be integrated.

¹⁰ <http://www.amazon.com/>

¹¹ <http://www.flickr.com/>

For example, Code Fragment 1 shows anonymised review and user data retrieved from the *Amazon Associates Web Service API*¹² in response to a `CustomerContentLookup` operation. The operation takes an *Amazon* `CustomerID` as input and in this case returns all information the customer has made public about themselves (the `CustomerFull` response group was requested).

`CustomerIDs` can be obtained by querying the API for reviews of a known item; these `CustomerIDs` can then be used in `CustomerContentLookup` operations to obtain additional data about the user. As Code Fragment 1 shows, personal information such as name, location and nickname are returned in such queries. Occasionally a user's email address is used as the value of the nickname field, however this is not consistent and in most cases no data is available that can be used to uniquely identify a user as a basis for integration with other types of data from external data sources, such as social network information.

¹² <http://aws.amazon.com/>

```

<Customer>
  <CustomerId>A89YPC0B3HML7X</CustomerId>
  <Nickname>joebloggs</Nickname>
  <WishListId>7DCW9CVSFW7RI</WishListId>
  <Location>
    <UserDefinedLocation>Bloggsville, Arizona, United
      States</UserDefinedLocation>
  </Location>
  <CustomerReviews>
    <TotalReviews>18</TotalReviews>
    <TotalReviewPages>2</TotalReviewPages>
    <Review>
      <ASIN>1234567890</ASIN>
      <Rating>4</Rating>
      <HelpfulVotes>2</HelpfulVotes>
      <Reviewer>
        <CustomerId>A89YPC0B3HML7X </CustomerId>
        <Name>Joe Bloggs</Name>
        <Location>Bloggsville, Arizona, United States</Location>
      </Reviewer>
      <TotalVotes>2</TotalVotes>
      <Date>1998-08-29</Date>
      <Summary>A great account of a tricky situation</Summary>
      <Content>In this witty book author Joe Bloggs recounts the
        challenges of being given one of the world's most common
        names. </Content>
    </Review>
  </CustomerReviews>
</Customer>

```

Code Fragment 1. Example of the structure of a CustomerFull Response Group to a CustomerContentLookup Operation on the Amazon Associates Web Service API.

Similarly, Code Fragment 2 shows (in *JSON* format¹³) the raw output of my network on *Del.icio.us*. In this case each user is identified simply by their *Del.icio.us* 'screenname', which allows further requests to be made to retrieve their bookmarks, but does not allow this data to be reliably integrated with that from other sources. As a result, the same individual may be listed as part of one's social network on *Del.icio.us* but not on another site, with no way of integrating this data across the two services. Ideally one would be able to port groups of contacts between services or maintain a central repository of social

¹³ <http://www.json.org/>

network information to which selected services were granted access, but this is not possible with current approaches.

```
[ "apathetic", "bertrand_sereno", "bjc", "captsolo", "castagna", "cayzers", "c  
onnolly", "csf", "cshirky", "cygri", "dctanner", "drewpca", "dublinclontarf",  
"earl", "elijah", "elked", "hartz", "hex", "holygoat", "hughglaser", "ianalche  
my", "inkel", "iredwards", "jesper", "jvw", "kasei", "kidehen", "lassila", "ld  
odds", "Leonya", "libby", "majorerror", "meisenstadt", "mischa_tuffield", "mo  
rtenf", "mpasin", "ndw", "nmg", "NormanP", "paddy", "psychemedia", "rogargon",  
"splintered", "stefano.bertolo", "t0m", "takeoneonion", "thomasfranz", "timo  
hannay", "tmcandrew", "wcrosbie", "xenic", "zephoria", "zool" ]
```

Code Fragment 2. The author's Del.icio.us social network as a JSON array, retrieved from the Del.icio.us API at <http://del.icio.us/feeds/json/network/tomheath>

These examples demonstrate that while Web2.0 APIs can help avoid the creation of 'data silos' or 'walled data gardens', the output from these APIs is not always easily integrated with that from other sources. This creates an additional problem where 'islands' of data are exposed to the Web but without links between related items in different data sets.

Addressing these issues requires the publishing of data in formats that are easily processed by third parties and that afford integration and interlinking with other data on the Web. Semantic Web technologies such as RDF (described in Section 5.3) and ontologies such as FOAF (Section 5.4) provide a solution to both these issues, as RDF affords easy reuse and linking while FOAF provides a common (and extensible) schema for describing people and social networks.

5.3. The Semantic Web

The Semantic Web (Berners-Lee, Hendler and Lassila, 2001, Shadbolt, Hall and Berners-Lee, 2006) takes core components of the Web architecture (Jacobs and Walsh, 2004) such as URIs and HTTP, and applies these to *data* as well as *documents*. Therefore whilst the conventional Web can be seen as a Web of *linked documents* (primarily HTML documents, but also images, movies etc.), the Semantic Web is a Web of

machine-readable, linked data (Berners-Lee, 2007). The result is a platform for large-scale integration of heterogeneous data, ultimately for the benefits of users (McBride, 2002).

5.3.1. Resource Description Framework (RDF)

Data on the Semantic Web is not published as tables or lists in HTML documents, but as '*triples*' according to the '*Resource Description Framework*' (Klyne and Carroll, 2004). RDF defines both a graph-based data model based on *subject, predicate, object* triples, and the RDF/XML format (Beckett, 2004) through which an RDF graph of one or more triples can be serialised as an XML document¹⁴. The subject of any RDF triple must be a URI or a 'blank node', the predicate must be a URI, and the object can be either a URI, a Literal or a blank node (Klyne and Carroll, 2004).

Publishing data in RDF conveys a number of benefits: data is machine-readable, easily integrated for querying or other forms of processing, and easily linked across disparate sources. Traditional data formats such as Comma Separated Variables (CSV), 'vanilla' XML and even HTML can all be described as machine-readable, as data can be represented in these formats and parsed reliably by software applications. However, data represented in RDF is machine-readable in a different way. Not only is it machine-readable at a syntactic level (i.e. it can be parsed reliably) but also at a semantic level, in that the meaning of RDF data is made explicit.

¹⁴ Several other serialisation formats are available, such as 'N-Triples', however RDF/XML is the officially recommended serialisation format. Note that the underlying semantic structure of an RDF graph remains constant regardless of the serialisation format.

The meaning of data described in RDF is indicated by the use of classes and properties (relations) taken from shared ontologies available on the Web and identified by a URI. For example, an RDF data publisher wishing to describe a person may choose to state in triples that a person *A* is an instance of the class '*Person*'¹⁵, defined in the 'Friend of a Friend' (FOAF) ontology (a popular vocabulary for describing some basic characteristics of people) (Brickley and Miller, 2007). The publisher may also wish to state that person *A* has the name 'Joe Bloggs' and the homepage `http://www.joebloggs.com/`. This is possible using the 'name' and 'homepage' properties of FOAF, as shown below in Code Fragment 3.

```
<?xml version='1.0' encoding='UTF-8'?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">

  <foaf:Person rdf:about="http://www.joebloggs.com/joe">
    <foaf:name>Joe Bloggs</foaf:name>
    <foaf:homepage rdf:resource="http://www.joebloggs.com/" />
  </foaf:Person>

</rdf:RDF>
```

Code Fragment 3. Example RDF document demonstrating the use of the `foaf:name` and `foaf:homepage` properties

5.3.2. Using RDF for Data Integration

Because classes and properties of ontologies on the Semantic Web are uniquely identified by URIs, multiple data publishers can reference the same elements in different locations thereby indicating that they subscribe to a shared definition of these terms. This ability to reference shared ontologies significantly streamlines the data integration process, as mappings do not need to be made between different ontologies when

¹⁵ `http://xmlns.com/foaf/spec/#term_Person`

integrating data from disparate sources. In reality two publishers may choose to describe their data using elements from different ontologies; this does make the data integration process slightly more complex, however mappings can easily be defined between classes or properties in different ontologies on the Web in order to address such situations.

Key to the flexibility of RDF for data integration is the ability to mix statements within one graph (which may be serialised as an RDF document) that use elements from any number of arbitrary ontologies, without the entire document needing to validate against a fixed schema. An RDF/XML document must simply be valid XML; there are no constraints on the statements made within the graph it serialises.

This contrasts with XML-based data interchange where all parties must agree on a common schema for documents, and arbitrary, heterogeneous data cannot be integrated if it does not conform to this schema. Therefore, integrating data from different sources may involve rewriting existing schemas if new information is to be incorporated and republished in the same document.

The consequences of this limitation is that XML-based data interchange and integration is often restricted to specific operations between partners in well-defined domains, or in the case of Web2.0 APIs XML data is integrated (at great cost in development effort) and republished on the HTML Web as 'mashups'. This approach does not scale well to large numbers of data sources for the following reasons: no common query language is implemented across Web2.0 APIs therefore specific code may need to be written to interact with each; all sources must expose data using a common XML schema or the programmer must transform data to their own abstract data model in order for it to be integrated; this integration generally takes place at the level of a relational database or a

data structure in memory thus limiting the ability to expose this data for reuse *as-is*; and lastly once integrated data is generally exposed as HTML or JavaScript-based mashups, thereby losing much of the semantics of the data and preventing its easy reuse.

5.3.3. Linked Data

Another difference between RDF and vanilla XML is that RDF allows machine-readable links to be created to other data. Whilst an XML Schema may define a `<uri></uri>` element to be populated with the URI of some item, the semantics of this relationship are not explicit. Consequently, and in contrast to RDF, machines cannot infer links between data based on such elements. This situation is analogous to enclosing a URL in `` tags within an HTML document (without using anchor tags ``) and expecting applications to interpret this string as a link.

A crucial feature of RDF is the ability to explicitly link data together across different sets of data on the Web. This is achieved by creating triples in which the subject and object are URIs from different data sources. The data publisher may choose which predicates from which ontologies are used in such 'RDF links' (Bizer, Cyganiak and Heath, 2007); some may specify more conventional properties of a resource, such as in Code Fragment 3, whereas others may state that two different URIs represent the same resource, as in Code Fragment 4. This highlights the fact that, just as regular Web pages can link to any other page, RDF statements can be made in any location on the Web referencing any URI, irrespective of the 'ownership' of the URIs being referenced.

```

<?xml version='1.0' encoding='UTF-8'?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/">

  <foaf:Person rdf:about="http://kmi.open.ac.uk/people/tom/">
    <foaf:name>Tom Heath</foaf:name>
    <owl:sameAs rdf:resource=
      "http://identifiers.kmi.open.ac.uk/people/tom-heath/" />
    <owl:sameAs rdf:resource="http://revyu.com/people/tom/" />
    <owl:sameAs rdf:resource=
      "http://my.opera.com/tomheath/xml/foaf/#me" />
  </foaf:Person>

</rdf:RDF>

```

Code Fragment 4. Example RDF document showing the use of `owl:sameAs` to state that a number of different URIs represent the same resource

Berners-Lee (2007) outlined four 'rules' which should be followed when producing 'Linked Data' for the Web:

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names
3. When someone looks up a URI, provide useful information
4. Include links to other URIs so that they can discover more things

5.3.4. Querying RDF Data

Web Services that publish vanilla XML require application developers to parse XML trees to retrieve the desired data. Whilst most programming languages provide libraries that make this task trivial, data processing remains tied to the underlying syntactic rather than semantic structure of the data, which may vary significantly across data sources. Creating Web2.0 mashups consequently requires the writing of custom handlers to

interact with each API. No common language is available for querying and integrating such data sources, and economies of scale through reuse of common schemas are rarely available.

Once RDF data has been integrated or linked in some fashion, the resulting graph can be queried using the SPARQL Query Language for RDF (Prud'hommeaux and Seaborne, 2007), the SQL-like query language being standardised through the W3C. SPARQL enables standardised access to distributed data sources. Queries are executed as HTTP GET requests against remote 'endpoints', returning data that can be processed using standard code, irrespective of the endpoints underlying implementation. Developers must simply know the structure of the RDF graph behind the endpoint in order to write the appropriate query.

5.3.5. Summary

These characteristics make RDF an ideal technology for flexible integration of heterogeneous data sources on the Web. In the context of this research, Semantic Web technologies support the integration of social network information with trust metrics, computed using evidence from multiple sources across the Web.

It should be noted that in this case Semantic Web technologies are used simply as a platform for integration of heterogeneous data. This is in contrast to work by authors such as Loizou and Dasmahapatra (2006) and Cantador, Castells and Bellogín (2007) that attempts to exploit ontologies as the basis for providing item-centric recommendations.

5.4. Sources of Social Network Data

Many potential sources of social network data exist on the Web, particularly in social networking sites such as *Facebook* and *LinkedIn*. However, as discussed above and despite the availability of the *Facebook Platform*¹⁶, the data held by these sites is not published in formats that afford easy integrating and linking with data from other sources. Consequently this research uses social network data published in RDF using the 'Friend of a Friend' (FOAF) vocabulary (Brickley and Miller, 2007).

Taking an RDF-based approach affords users greater choice and flexibility in how their personal information is managed and published, as data can be made available in locations of their choosing and under their control, from where it can be shared with third party applications.

The FOAF vocabulary provides properties and classes for describing common features of people and their social networks. The basic unit for defining social relationships in FOAF is the `knows` property, simply used to state that Person A knows Person B. This degree of semantics is sufficient for many application scenarios, and avoids potentially awkward social situations arising from individuals having different perceptions of the nature of a relationship.

Other vocabularies, such as the Relationship vocabulary (Davis and Vitiello, 2005), have been proposed that go beyond the shallow semantics of `foaf:knows` to describe greater subtleties in the relationships between individuals. The greater specificity provided by

¹⁶ <http://developers.facebook.com/>

such vocabularies may be beneficial for certain applications, but is unlikely to enhance this research as it is not apparent how different relationship types may predict trust relationships between individuals in the domains with which this research is concerned.

For example, one could predict that in general a spouse or partner would be trusted to a greater extent than an acquaintance. This may be the case at a general level; however, in an information-seeking scenario an acquaintance with relevant domain knowledge may be more highly trusted than a spouse as a source of information in that domain.

Ashri, Ramchurn, Sabater et al. (2005) describe an approach that exploits the nature of social relations to determine the trustworthiness of other agents in multi-agent systems. However, due to the domain of their work the relationships used are heavily market-oriented, such as 'trade', 'dependency', 'competition' and 'collaboration' and consequently are not applicable to this research.

For these reasons, and as the study reported in Chapter 4 did not identify any specific effects of the type of relationship on source selection in recommendation-seeking, the `foaf:knows` relationship is deemed adequate as a definition of social relations in this research.

Ding, Finin and Joshi (2005) report that there are nearly one million instances of the `foaf:Person` class on the web, distributed among roughly 45,000 documents. A number of services such as *LiveJournal*¹⁷, *Tribe.net*¹⁸ and *MyOpera*¹⁹ do publish FOAF data about

¹⁷ <http://www.livejournal.com/>

¹⁸ <http://tribe.net/>

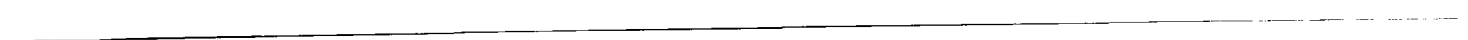
their registered users, whilst many members of the Semantic Web community maintain their own FOAF files on their personal Web sites. How FOAF data is harvested from the Web and used to support this research is described in Section 7.5.6.

5.5. *Implemented Systems and Detailed Architecture*

For the reasons outlined above, the systems developed in the course of this research are distributed in nature and based on Semantic Web technologies. Two user-oriented systems have been developed: *Revyu.com* and *Hoonoh.com* (hereafter simply Revyu and Hoonoh). These are complemented by two systems not designed for human users: a *Trust Computation Subsystem* and a *FOAF data repository*.

Revyu is a Web site that allows people to create reviews and ratings that are then published in RDF on the Semantic Web for reuse by other applications. By making review data available in RDF, Revyu overcomes the limitations of review data provided by services as *Amazon*, as discussed above. This review data provides input to the Trust Computation Subsystem, which along with data from the FOAF repository, powers Hoonoh. Hoonoh is a Web site that provides source-centric information-seeking on the Web supported by trusted social networks.

Figure 5 below shows the system architecture in detail.



¹⁹ <http://my.opera.com/>

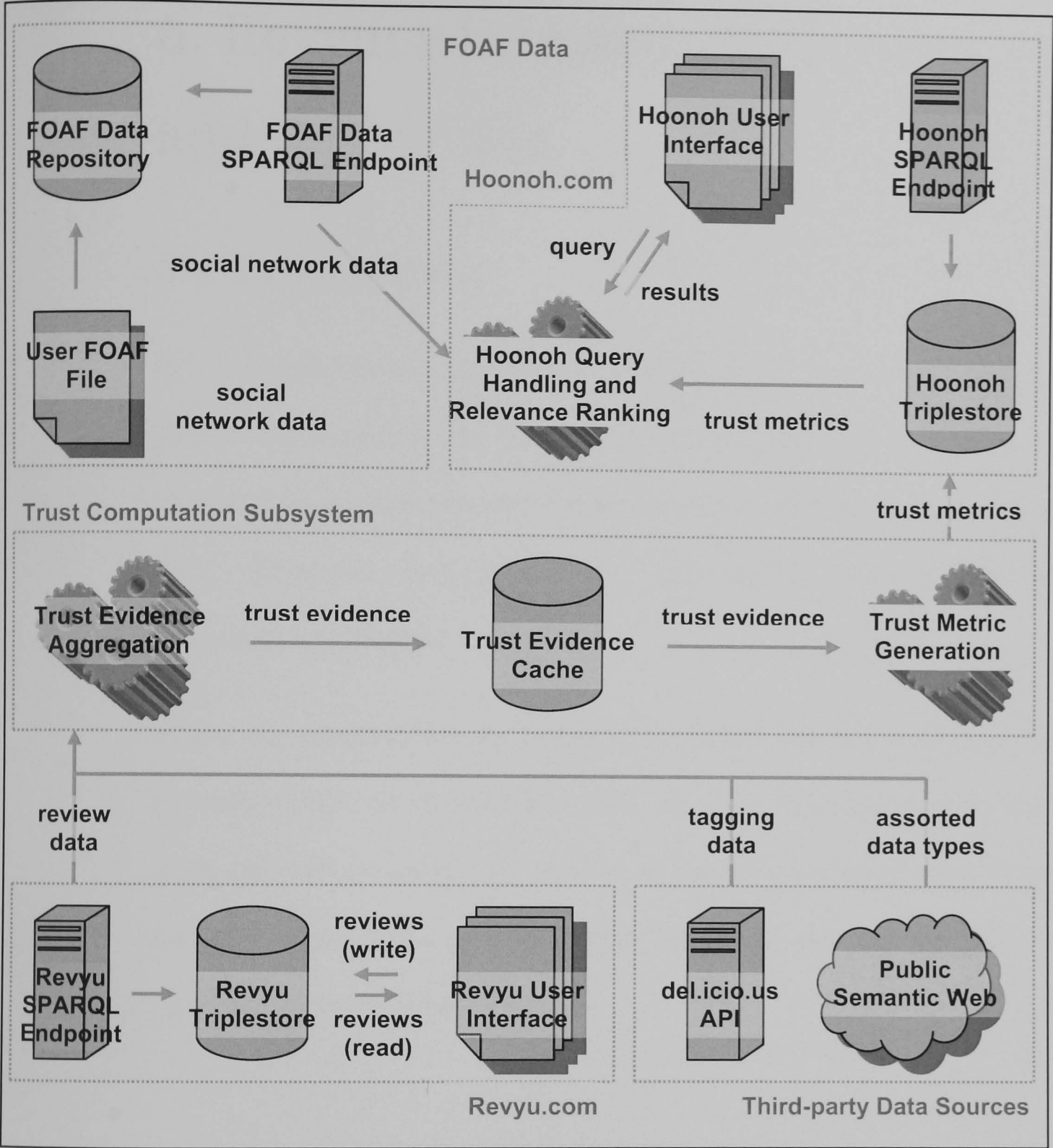


Figure 5. Architecture of systems developed to support, or used within, this research

Revyu is presented in detail in the next chapter. Following that, Chapter 7 describes: the 'Hoonoh algorithms' for generating trust relationship metrics based on data from Revyu and other sources on the Web, the broader 'Trust Computation Subsystem' in which these algorithms are instantiated, and the Hoonoh system itself. The trust metrics on which Hoonoh is based are also published on the Web in RDF to enable reuse by other applications; see Section 7.4 for details.

6. Revyu: a Semantic Web Reviewing and Rating Site

6.1. Introduction

Revyu was developed to enable the collection of data from which trust metrics could be derived and integrated with social network data. Revyu is a reviewing and rating site in the mould often associated with Web2.0 but which has been present on the Web for some time²⁰. Prominent examples of such sites include *Epinions* and the reviewing functionality of *Amazon*.

Revyu was launched as a live, publicly accessible Web site at <http://revyu.com/> in November 2006. As of November 2007 more than 650 reviews have been created by more than 150 reviewers. The reviews in the system cover a range of types of items including books, films, concerts, hotels, restaurants and academic papers. The Revyu homepage is shown in Figure 6 below.

²⁰ The Wayback Machine at http://web.archive.org/web/*/http://www.epinions.com/ indicates that Epinions has existed since at least 1999.

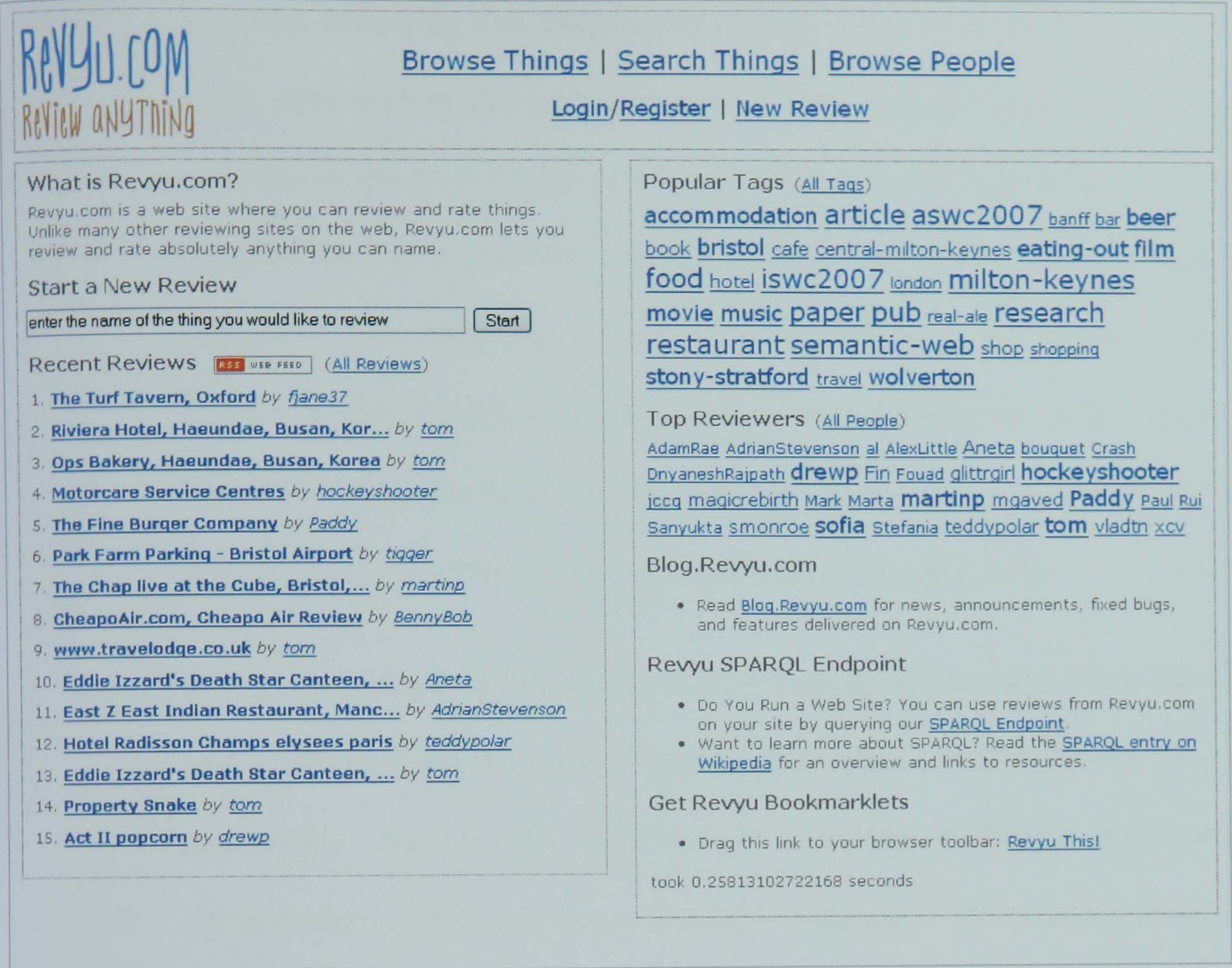


Figure 6. The Revyu Home Page

6.2. Novel Features of Revyu

Revyu differs from existing Web-based review and rating systems in a number of significant ways. Firstly, users of the site are not restricted by the closed worlds of conventional reviewing sites that limit reviews to items from a specific domain, sold by a particular company, or catalogued in an existing database. Instead Revyu takes a more open-world approach where users are free to review anything they choose. In addition to giving the user flexibility this has the benefit of not requiring a database to be maintained of items suitable for review, as is the case with existing cross-domain review sites such as *Epinions*.

Secondly, reviewing sites that provide data for reuse via an API are not widespread. As a result, sites such as *Epinions* and *TripAdvisor*²¹ become closed world silos of reviews available on the Web but not well interlinked with other relevant data. Even where APIs are provided, by *Amazon* for example, these reviews are generally made available in formats such as XML that do not afford interlinking at the data level, as discussed in Chapter 5. This hinders the interlinking and aggregation of all reviews of a particular item from across the Web, because without the use of universal identifiers such as URIs it is not easy to determine if two reviews refer to the same item.

To overcome these issues, Revyu is built natively on Semantic Web technologies. As a result, the site identifies reviews (and all other types of objects in the system) with URIs and exposes these on the Web in RDF according to the principles of Linked Data (Berners-Lee, 2007), and via a SPARQL endpoint²². This enables reuse of data from Revyu in third party applications, more flexible querying via SPARQL, and easier integration and linking of data across different sources, as previously outlined in Chapter 5.

Thirdly, Revyu exploits this ease of data integration to enhance the site with data from external sources without requiring this data to be replicated at Revyu.

Lastly, the majority of conventional reviewing and rating sites only identify reviewers by nicknames or unique identifiers that have only local rather than global scope. As a result one can rarely base decisions about the trustworthiness or value of a review on pre-

²¹ <http://www.tripadvisor.com/>

²² <http://revyu.com/sparql>

existing knowledge of the reviewers, as nicknames obscure their true identity and prevent one from identifying all reviews by known and trusted individuals. Instead, characteristics such as writing style must be relied upon in judging the suitability or trustworthiness of a review.

To overcome this and enable integration of reviews with social network data Revyu includes a SHA1 hash (Eastlake and Jones, 2001) of the reviewer's mailbox URI in its RDF output of reviews, using the `mbox_sha1sum` property from the FOAF vocabulary²³. This serves to uniquely identify a reviewer without disclosing his identity to those who do not already know his email address, as the SHA1 algorithm makes it "computationally infeasible to find a message which corresponds to a given message digest, or to find two different messages which produce the same message digest" (Eastlake and Jones, 2001) (pp. 2).

After briefly highlighting some related work, the remainder of this chapter presents Revyu from a user and implementation perspective and discusses design decisions made during creation of the site.

6.3. Related Work

The idea of using RDF to publish reviews on the Web is not new. Golbeck and Hendler (2006) expose film reviews in RDF via the FilmTrust system. Revyu improves upon the functionality offered by FilmTrust, as users of that system are restricted to reviewing and

²³ http://xmlns.com/foaf/spec/#term_mbox_sha1sum

rating items in just one domain (films), reviewed films are not annotated in any way beyond the rating, and the accumulated ratings can not be queried programmatically.

Revyu takes a significant and concrete step beyond this by exposing reviews via RDF and SPARQL according to the Linked Data principles discussed in Chapter 5. In doing so it creates a major node in a potentially Web-wide ecosystem of interlinked reviews and ratings, and helps to bootstrap the Semantic Web as a whole.

Revyu goes beyond the work of Guha (2004) by implementing an open rating system that supports the reviewing and rating of anything, not just Web content. Furthermore, the trust metrics developed in this research, based on the study reported in Chapter 4, are more fine-grained than Guha's trust/distrust distinction, more task- and context-sensitive and are computed automatically without relying on manual ratings of others in the network. The algorithms by which these trust metrics are calculated will be described in Chapter 7.

As an application that generates RDF data from user input, Revyu warrants some comparison to generic semantic annotation mechanisms such as Semantic Mediawiki (Völkel, Krötzsch, Vrandečić et al., 2006). This extension for the popular MediaWiki²⁴ wiki engine has generated considerable interest and gained some noteworthy uptake in sites such as DiscourseDB²⁵, however it is not apparent whether the application is sufficiently usable or compelling to elicit semantic annotations from non-specialists.


²⁴ <http://www.mediawiki.org/>

²⁵ <http://discoursedb.org/>

Conversely, applications exist that allow users to create arbitrary, ontology-based, annotations of a specific type of object. PhotoStuff (Halaschek-Wiener, Golbeck, Schain et al., 2005) is a desktop application that enables ontology-driven semantic annotation of photographs. However, it is not clear whether requiring users to annotate photos with elements taken directly from specifically-loaded ontologies will scale to annotations across a wide range of domains, and whether installation of a desktop application hinders uptake compared to tagging-based photo annotation applications such as *Flickr*.

6.4. User Walkthrough

Users can search or browse the site to read existing reviews, descriptions of things reviewed on the site, and profiles of reviewers. To the non-specialist Revyu appears like any regular Web site: little indication is given that it is based on Semantic Web technologies. All site content is published in HTML and RDF/XML, however users viewing the site with a conventional Web browser will never be exposed to the underlying RDF data unless they explicitly request it, either by clicking a link in HTML pages on the site or by sending appropriate `Accept` headers in their HTTP request (as discussed below in Section 6.9). Figure 7 shows a review created on Revyu, as it appears in a conventional Web browser.



[Home](#) | [Browse Things](#) | [Search Things](#) | [Browse People](#)
[Login/Register](#) | [New Review](#)

Review of: The Fine Burger Company

★★★★☆ by [Paddy](#) on 05 Nov 2007

I recently visited my local branch of this relatively new chain. Burgers (beef, lamb, chicken, veggie) were the only main meal choice on the menu although you could customise your burger according to your fancy. Prices first appeared acceptable for a decent burger (see their website) but the cost of accompaniments (e.g chips @ £2.70) quickly pushed the prices up.

The staff were very friendly, perhaps in need of better organisation/training – we saw nobody for a while then were asked if we were ready to order 4 times in about 90 seconds. No big deal though. The surroundings were pleasant. It was quiet for a Saturday evening.

The food was delicious, it was indeed a fine burger and the chips were also tasty although I've had much better. The salad that we had ordered didn't arrive.

When it was time to leave, I asked for the bill and mentioned the salad hadn't arrived and requested that they check it wasn't on the bill. The bill arrived, the salad wasn't included but a 12.5% Service Charge had been added. The service wasn't great, especially since the salad was totally forgotten about. Soft drink re-fills were self service, there were 3 of us and we'd only had one course. Stealthily adding a service charge seemed outrageous. I politely asked for that to be removed too, pointing out the missing salad and that I'd like to have the decision whether to tip and if so, by how much. The waitress was somewhat taken aback at my objection and said she'd need to call the manager. Which she did and it was duly removed.

I left a £2 cash tip (5% of the bill) and we left. I thought this was generous but served to amplify the case in point.

Summary: good food, limited and perhaps a little pricey but a Service Charge policy that stinks.


Review Permalink: <http://revvu.com/reviews/bdefdf8390dafe4e8d33fed7492eabce1a4f1df>

The Fine Burger Company

Tags: [burger](#) [dining-out](#) [fine-burger-company](#) [restaurants](#) [service-charge](#) [tip](#)

Homepage: <http://www.fineburger.co.uk/>

What do you think of The Fine Burger Company? [Write Your Own Review...](#)

[Write a Review of The Fine Burger Company](#)
[RDF Metadata for this Review of The Fine Burger Company](#)

[Add to del.icio.us](#)

[Revyu.com](#): [Contact](#) | [Credits](#) | [Privacy Policy](#) | [Disclaimer](#)

Figure 7. The HTML view of a review on Revyu

6.4.1. Generating Semantic Web Content by Completing Web Forms

Users who wish to create reviews and ratings can do so simply by registering with the site and filling in a Web form as shown in Figure 8. The reviewing form can be accessed by following a link on the Revyu site or using the Revyu 'bookmarklet', a browser widget that redirects the user from the site they are currently viewing to the reviewing page on Revyu; this can be helpful where the user wants to review a certain Web page or a thing described by the Web page, as a relationship between the reviewed item and the origin Web page is recorded by Revyu (see Section 6.7).

Revyu.com

Review anything

[Home](#) | [Browse Things](#) | [Search Things](#) | [My Network](#) | [All People](#)

Hello Tom Heath! | [My Account](#) | [Logout](#) | [New Review](#)

Review Something

You Are Reviewing...

Is this a good name for the thing you are reviewing? If not, please make any changes now.

Your Comments...

Letters, number, spaces, and punctuation are permitted. Sorry, no HTML tags.

Your Rating...

No Rating

Please provide a rating

Tags...

add some tags

Enter some keywords that describe
Separate keywords with spaces.
If you need to join words, use a hyphen
e.g.: tourist-attraction museum new-york

Figure 8. The upper half of the Revyu Reviewing Form

The Revyu reviewing form in Figure 8 simply asks users to provide a name for the thing they wish to review, the text of their review, a numerical rating (on a scale of 1-5, where 1 represents Very Bad, and 5 represents Very Good), some keyword tags related to the thing being reviewed, and one or more links to related Web resources.

This mode of interacting will be familiar to those who have written reviews at sites such as *Epinions* or *Amazon*, and is designed to enable novice users to contribute reviews through a Web2.0-style interface, but make these reviews available online in the appropriate Semantic Web format.

Web2.0 applications and services such as *Wikipedia*²⁶, *Flickr* and *Del.icio.us* have enabled non-specialist users to contribute to the Web on a scale that is inline with the original vision of a 'read-write Web' (Berners-Lee and Fischetti, 2000), but had not previously been achieved. This has been made possible by providing simple, well-structured interfaces based on Web forms, through which users can, for example, edit wiki entries or tag photos and bookmarks. Such interfaces lower the cost of adding content and annotations to the Web compared to traditional publishing techniques that involve specialist skills and software.

Following a similar approach, Revyu is designed to be usable by humans whilst transparently generating machine-readable RDF metadata based on their input. By adhering to this well established interaction pattern, Revyu allows users to create Semantic Web data that can be used in computing trust metrics for this research, without requiring any knowledge of RDF.

In an evaluation of Semantic Web applications deployed to members of the Semantic Web community (Heath, Domingue and Shabajee, 2006) it was found that the usability of applications hindered their uptake, even by those knowledgeable in the field. In the light of these findings, tools that make semantic annotation feasible for specialists and non-specialists alike are required if user-generated Semantic Web data is to be created on a significant scale.

To date users of Revyu have created over 20,000 RDF triples which are publicly available on the Semantic Web. Whilst not a large figure by some standards, it is

²⁶ <http://www.wikipedia.org/>

significant that these triples have been generated primarily from direct user input, rather than by data mining or extraction from natural language.

Reviews submitted through the reviewing form are converted to RDF and stored as persistent triples in the Revyu triplestore (see section 6.8 below). From there they are immediately available on the site in HTML and RDF/XML, and via the Revyu SPARQL endpoint.

6.5. *Use of Common Ontologies*

Revyu uses the FOAF ontology (Brickley and Miller, 2007) to describe reviewers. As discussed previously, the FOAF ontology includes a property 'mbox_sha1sum', the value of which allows reviewers to be uniquely identified while only making their identity visible to others who already know their email address.

Review data is published using the Review RDF vocabulary (Ayers and Heath, 2006), which has properties for describing aspects of reviews crucial for calculating trust metrics in social networks, such as numerical ratings and the creator of a review. The RDF output of a review on Revyu is shown in Figure 9.


```
<?xml version="1.0" encoding="UTF-8" ?>
<rdf:RDF
  xmlns:base="http://revyu.com/"
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:dc="http://purl.org/dc/elements/1.1/"
  xmlns:dcterms="http://purl.org/dc/terms/"
  xmlns:vcard="http://www.w3.org/2001/vcard-rdf/3.0#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:rev="http://purl.org/stuff/rev#"
  xmlns:tag="http://www.holygoat.co.uk/owl/redwood/0.1/tags/"

  <rdf:Description rdf:about="things/motorcare-service-centres-tyres-exhausts-brakes-batteries">
    <rdfs:label>Motorcare Service Centres</rdfs:label>
    <rev:hasReview rdf:resource="reviews/ecb44c5cb5b386ca8caa664546b786badee0a8cd"/>
  </rdf:Description>

  <rdf:Description rdf:about="reviews/ecb44c5cb5b386ca8caa664546b786badee0a8cd">
    <rev:createdOn rdf:datatype="http://www.w3.org/2001/XMLSchema#dateTime">2007-11-10T04:43:15-08:00</rev:createdOn>
    <rev:reviewer rdf:resource="people/hockeyshooter"/>
    <rev:maxRating rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">5</rev:maxRating>
    <rev:minRating rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">1</rev:minRating>
    <rev:rating rdf:datatype="http://www.w3.org/2001/XMLSchema#integer">5</rev:rating>
    <rev:text>The particular garage I visited is one of a chain of three. I have been there once before, and a friend
  </rdf:Description>

  <rdf:Description rdf:about="things/motorcare-service-centres-tyres-exhausts-brakes-batteries">
    <rdfs:seeAlso rdf:resource="things/motorcare-service-centres-tyres-exhausts-brakes-batteries"/>
  </rdf:Description>

  <rdf:Description rdf:about="people/hockeyshooter">
    <rdfs:seeAlso rdf:resource="people/hockeyshooter"/>
  </rdf:Description>

  <rdf:Description rdf:about="reviews/ecb44c5cb5b386ca8caa664546b786badee0a8cd">
    <rdfs:label>Review of Motorcare Service Centres, by hockeyshooter</rdfs:label>
    <foaf:primaryTopic rdf:resource="things/motorcare-service-centres-tyres-exhausts-brakes-batteries"/>
  </rdf:Description>

</rdf:RDF>
```

Figure 9. Example of RDF output of a Revyu review

The Tag ontology (Newman, Russell and Ayers, 2005) is used to describe bundles of tags associated with reviewed items, when they were added, and by whom. This makes tagging data readily available for use in other applications, and in tag-interoperability initiatives such as the TagCommons (Gruber, 2007). Common properties from RDF, RDFS and OWL (such as `rdf:type`, `rdfs:label`, and `owl:sameAs`) are also used frequently within the RDF published by Revyu.

Adopting these popular ontologies makes Revyu data instantly interoperable with that from other sources. Creating a Revyu-specific ontology that was then mapped to others would have been an equally valid, albeit more complex process, that would have brought few benefits.

Revyu also exposes reviews using the hReview 'microformat'²⁷ embedded in XHTML pages. This makes Revyu content accessible to applications that currently support microformats but not RDF. Whilst popular among sections of the Web2.0 community, microformats do not provide the same data integration and linking capabilities of RDF.

6.6. The Role of Tagging in Revyu

6.6.1. Tagging versus Classification

A decision was made when designing and implementing Revyu to not require users to classify reviewed items according to an existing taxonomy, but instead allow them to tag with one or more descriptive keywords an item being reviewed.

This decision was made for both user-oriented and implementation-related reasons: classifying reviewed items would require the user to identify an appropriate category in an existing, fixed taxonomy to which not all reviewers could subscribe. Furthermore, if users were to be given complete flexibility in what they reviewed then such a classification would by definition be large and therefore complex. A sufficiently comprehensive classification was not readily available, and even the entire range of ontologies available on the Web were not seen to provide adequate coverage of all types of items that users might wish to review. Even were this was not the case, developing a sufficiently usable interface with which users could easily categorise any item was considered unfeasible.

²⁷ <http://microformats.org/wiki/hreview>

As a result, keyword tagging was chosen in favour of classification, as this can aid other users of the site in browsing or searching for reviews, whilst not creating barriers to the contribution of reviews and allowing for reviewing of items that might not be easily categorised but can be described with a few keywords.

When users start entering tags in the Tags field of the Revyu reviewing form, suggestions are displayed of tags they may want to use based on those already present in the system. This helps avoid spelling mistakes, aids convergence on particular syntactic forms, and ensures consistency of tag usage.

A less desirable consequence of the use of tagging in Revyu is that machine-readable statements regarding the nature of reviewed items cannot be made with any confidence from tagging data alone. For example, the tag `book` may refer to a volume of reading material but to a service for booking concert tickets. Similarly, an item tagged `film` may not be a movie film but a particular brand of photographic film. Therefore, by default Revyu makes no assumptions about the type of reviewed items based on how they have been tagged.

By allowing less structured input from users the burden of identifying the 'type' of reviewed items is transferred to Revyu if the site is to provide additional functionality based on this information. Derivation of type information from tagging data is currently undertaken in two domains, books and films, using external data sources to help ensure accurate results. Similar heuristics may feasibly be implemented for items such as music albums, pubs, restaurants and hotels.

6.6.2. Inferring the Type of Reviewed Items

Identifying Films on Revyu

The majority of contemporary films have homepages, which are generally provided by the film studio but carry little if any machine-readable data about the picture. However, coverage of films is very high in *Wikipedia*, which provides an external source against which Revyu data can be verified by querying the *DBpedia* (Auer, Bizer, Lehmann et al., 2007) SPARQL endpoint²⁸. The following heuristic is used to identify films: for each reviewed item tagged 'film' or 'movie', look for items in *DBpedia* of type 'film' and with the same name. For any items for which this heuristic returns a match, an `rdf:type` statement is added to the Revyu triplestore asserting that this item is a film. This type information is exposed in the RDF descriptions of items on the Revyu site and also used to trigger retrieval of additional information about the reviewed item for display on the site, as described below in Section 6.11.

Identifying Books on Revyu

Whilst *Wikipedia* (and thus *DBpedia*) has extensive coverage of films, the coverage of books is less comprehensive; therefore a different heuristic is used to identify books reviewed on Revyu. When reviewing books, reviewers often place links to an *Amazon* Web page about the book in one of the Links fields of the reviewing form (generally the 'Other Links' field, as described below).

²⁸ <http://dbpedia.org/sparql>

Where these links exist they are parsed and analysed to extract ISBN numbers. If a valid ISBN is identified then an `rdf:type` statement is added to the Revyu triplestore asserting that this item is a book. Again, this type of information is used to retrieve additional information about the item, also as described below. Parsing links to external resources in this way is preferred over simply looking up all items tagged 'book', due to the potential for books and other items with the same name to cause false positives.

6.6.3. Identifying Related Tags

Many tags are used together when reviewing items, presumably because they are related in some way. An algorithm is used to identify tags that frequently co-occur (above a certain threshold of co-occurrence, to avoid identifying spurious connections) from tagging data in Revyu. For example, the algorithm finds that 'pub' is related to 'beer' and 'food'.

These relations are then logged in the Revyu triplestore and republished in both HTML and RDF. In the HTML pages about each tag²⁹, tags that co-occur above a certain threshold are displayed to the user. This threshold is set low for HTML output, as human readers of the page are unlikely to infer erroneous information based on these relationships. The RDF output uses the `skos:related` property of the SKOS vocabulary (Miles and Brickley, 2005), asserting that these two concepts are related. This makes these conceptual relationships accessible to other applications wishing to find information about connections between tags. In contrast to the HTML output,

²⁹ e.g. <http://revyu.com/tags/pub/about.html>

relationships exposed in RDF descriptions of tags are based on a more conservative threshold, in order to avoid erroneous inferences based on these assertions.

Finding co-occurrence relationships between tags is certainly not unique to Revyu; what makes this work more noteworthy is the republishing of these relationships to the Web in RDF. At present no attempt is made to link tags to other concepts in e.g. WordNet (van Assem, Gangemi and Schreiber, 2006), as sufficient accuracy can not be guaranteed, especially when dealing with homonyms. However, techniques described by Specia and Motta (2007) suggest how Revyu tags may be better integrated with the Semantic Web.

6.7. The Role of Links in Identifying Reviewed Items

As discussed earlier in this chapter, Revyu takes an open world view of the reviewing process by not constraining users to reviewing items from a fixed database; anything that the user can name can be reviewed. This has the potential to create a situation where an item has been reviewed, but the exact 'identity' of the item is not apparent from the content of the review. To minimise the occurrence of such situations Revyu allows reviewers to specify a number of links that are associated with the item being reviewed in one of three ways: the home page of the item, a page that contains additional information about the item but is not the home page, or the actual location of the item where it exists on the Web. Figure 10 below shows the three Link fields on the Revyu reviewing form.

Home Page...
The official home page of is at...

http://

This Home Page should be unique to the thing you are reviewing. If it isn't, please use the See Also field instead.

Other Links...
There is also information about at...

http://

Web-only Resource...
is something that only exists on the web, such as an online tutorial or a web page. It can be found at...

http://

Figure 10. The lower half of the Revyu reviewing form

These external links provide a way for human users of the site to disambiguate reviewed items in cases where there is any ambiguity. Disambiguation can also be carried out by applications that use Revyu's machine-readable RDF output, as the contents of these fields are saved as RDF triples when the review is submitted, using the `foaf:homepage`, `rdfs:seeAlso` and `owl:sameAs` predicates respectively.

The `owl:sameAs` property indicates that two URIs identify the same item, thereby linking a thing's representation on Revyu to its true location on the Web. RDF-aware users can also enter URIs that represent things other than Web documents ('non-information resources') into this 'Location' field in order to link Revyu-generated URIs to equivalent URIs minted by other data providers.

Links made using `rdfs:seeAlso` are of less value for these purposes, however the `homepage` property is defined in the FOAF ontology as 'Inverse Functional', meaning that the object of a `foaf:homepage` triple uniquely identifies the subject of the triple. Consequently it can be inferred that two resources that have the same `foaf:homepage` are in fact the same resource. This feature opens up the possibility of using Semantic

Web lookup services such as *Sindice* (Tummarello, Oren and Delbru, 2007) to identify other sources of information about items reviewed on Revyu.

6.8. Technical Implementation

Revyu is built from the ground upwards on Semantic Web technologies. This section describes the Revyu architecture and discusses decisions made in implementing the system.

The site is implemented as a Web application written in PHP³⁰ and running on a regular Apache Web server³¹. The creation, storage, querying, manipulation and publication of RDF data is supported by RAP, the RDF API for PHP (Oldakowski, Bizer and Westphal, 2005). RAP is a PHP library that provides programmatic methods for common RDF-related tasks. Apache *mod_rewrite* rules are used to provide 'pretty URIs' such as <http://revyu.com/people/tom> and to abstract the structure of URIs away from the details of the underlying implementation.

Upon completion and submission of the Revyu reviewing form the review, and all related information such as tags and Web links associated with the reviewed item, is converted into RDF triples and persisted to the Revyu triplestore. This triplestore is simply based on a de-normalised MySQL database structured according to the RAP database schema³².

³⁰ <http://www.php.net/>

³¹ <http://httpd.apache.org/>

³² http://sites.wiwiss.fu-berlin.de/suhl/bizer/rdffapi/database_schema.html

6.9. *Linked Data Compliance*

From the outset Revyu was designed to adhere to the rules of Linked Data (Berners-Lee, 2007) outlined in Chapter 5, which ensures that reviews hosted on the site can be fully connected into the Semantic Web. This section details how Revyu adheres to the rules of Linked Data.

Firstly, all entities on Revyu are given URIs: reviewed items (referred to simply as 'things'), reviews, reviewers, tags and even the bundles that represent tags assigned by one person at one point in time (known as 'taggings'). This enables linking between Revyu and other data sets on the Web. All URIs are HTTP URIs, with the base `http://revyu.com/`.

If the thing being reviewed has already been reviewed at Revyu then the existing URI for that thing is used as the subject of the review, otherwise a new URI is minted to identify this thing. URIs for things are created in the `http://revyu.com/things/` address space and based on a combination of the name given to the item by the reviewer and, in the case of short titles, tags associated with the item. Where this data alone would not yield a unique URI an additional timestamp is added to the URI. An example of a 'thing' URI in Revyu is shown in³³.

Reviewer URIs are minted based on the 'screenname' the reviewer chose when they registered with the site and take the form `http://revyu.com/people/screenname`, whilst reviews themselves are minted URIs in the `http://revyu.com/reviews/` address

³³ `http://revyu.com/things/designing-with-web-standards-by-jeffrey-zeldman`

space based on a SHA1 hash of the name of the reviewed item and the timestamp of the review. A bundle of tags (collectively referred to as a 'tagging') is identified by a URI based on the same hash used to generate the review URI, whilst URIs for keyword tags are simply based on the tag itself and take the form `http://revyu.com/tags/tag`. One consequence of minting URIs for tags based on their syntactic form is that homonyms cannot be distinguished; however this is a limitation of tagging in general, irrespective of whether or not URIs are minted for tags.

Providing URIs for all reviewed things gives many non-Web resources a presence on the Semantic Web which they would not otherwise have. This enables any third party to make reference to these items in other RDF statements without having to mint additional URIs, and is particularly useful for items, such as restaurants or pubs, that are unlikely to mint their own URIs in the near future. Consequently Revyu fulfils a valuable role in bringing new items into the Web of data.

Secondly, all URIs in the Revyu address space can be 'dereferenced'. Attempts to dereference the URIs of things which are not 'information resources' receive an HTTP303 'See Other' response, along with the URI of a document containing a description of the resource; this is commonly known as a '303 redirect'. The precise document to which the 'useragent' is redirected depends on the preferences for different types of content specified in the `Accept` header of the initial HTTP request. This is known as 'content negotiation' (Fielding, Gettys, Mogul et al., 1999), and allows conventional Web browsers to dereference URIs for non-information resources and be redirected to HTML documents that describe the resource, whereas Semantic Web browsers or other Semantic Web applications can be redirected to RDF descriptions of the same resource.

For example, a standard installation of *Mozilla Firefox*³⁴ sends `Accept` headers with the following value:

```
text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,text/plain;q=0.8,image/png,*/*;q=0.5
```

This indicates that `text/xml`, `application/xml`, `application/xhtml+xml` and `image/png` are the preferred media types, but if these are not available then the server should send `text/html`, followed by `text/plain`, followed by any other content type.

In contrast, a Semantic Web application may send an `Accept` header such as the following, indicating a preference for `application/rdf+xml` content:

```
application/rdf+xml, application/xml; q=0.9, text/xml; q=0.5
```

Content negotiation on Revyu URIs that represent non-information resources is carried out by a PHP script that analyses the `Accept` header of the request and redirects the useragent to either HTML or RDF documents that describe the resource. This configuration is inline with the W3C Technical Architecture Group's finding on the `httpRange-14` issue (W3C Technical Architecture Group (TAG), 2005), and serves to reinforce the distinction between a resource and a description of that resource.

³⁴ <http://www.mozilla.com/>

6.10. *Links to other Data Sets*

Where possible, links are made between Revyu data and items in external data sets (see Figure 11) in order to avoid Revyu data becoming an isolated island of RDF. Publishing these links in RDF connects Revyu in to a growing Web of Linked Data that is signified in particular by initiatives such as the Linking Open Data community project (Bizer, Heath, Ayers et al., 2007).

Many of these links are created during the same processes described above that attempt to derive type information from tagging data by validating against external sources. For example, where a reviewed film or book is found to exist in *DBpedia* or the *RDF Book Mashup* (Bizer, Cyganiak and Gauss, 2007), `owl:sameAs` statements are added to the Revyu triplestore to record that both URIs identify the same item. Likewise, where a user provides the URI of their FOAF file at registration time, `owl:sameAs` statements are made between the reviewer's Revyu URI and the URI they use to identify themselves in their FOAF description. These statements are then republished in the reviewer's RDF description on Revyu.

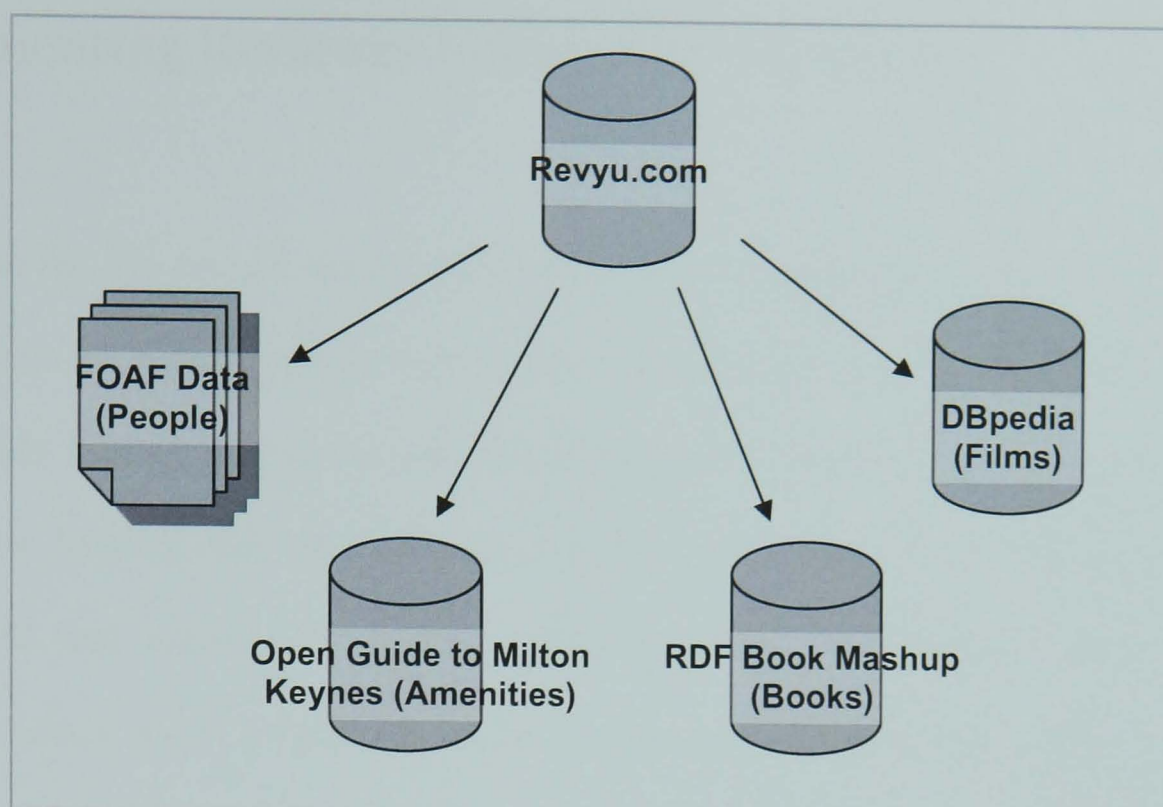



Figure 11. Links from Revyu to external data sets

6.11. Consuming Linked Data

Links between Revyu and external data sources are used as the basis for retrieving additional information about reviewed items from external Semantic Web data sources, without requiring the reviewer to provide this information. This information is shown alongside review data from Revyu in the HTML pages about an item, thereby enhancing the experience provided to users of the site without placing an additional burden on reviewers. The following sections provide details of how this is carried out. It is worth noting that a slightly different approach is taken for RDF documents describing items on Revyu. In this case `owl:sameAs` links between items are exposed but without republishing RDF data from external sources. The rationale for this is that 'true' Semantic Web browsers are not expected to be document browsers but data browsers. Therefore such applications will need to aggregate information from numerous sources before presenting a composite view to the user, in which case republishing third party data at Revyu would simply lead to unnecessary duplication.

6.11.1. Supplementing Reviewer Information with FOAF Data

Users registering with the site are not asked to provide copious information to populate their user profile, only an email address, screenname and password (real name can optionally be provided). Instead, where a reviewer maintains their own RDF (i.e. FOAF) description in another location they may also provide its URL. In this case Revyu dereferences this URI and queries the resulting graph for relevant information the reviewer chooses to share about themselves, such as photographs, homepage links, interests, and locations. This information is then used to enhance the reviewer's (HTML) profile page (as illustrated in Figure 12), thereby exploiting the data integration capabilities of a Semantic Web to provide the kind of rich user profiles often associated with Web2.0 applications without the information needing to be duplicated in Revyu. This approach reduces the burden on the user by not requiring them to manage multiple redundant sets of personal information stored in different locations, as one central set of personal information can be maintained in their FOAF file.



[Home](#) | [Browse Things](#) | [Search Things](#) | [Browse People](#)
[Login/Register](#) | [New Review](#)

Reviews by tom (176)

[Korea, Lonely Planet Country Guide, by Rob Whyte](#)

★★★★☆ on 22 Nov 2007

Using this guide book confirmed by growing suspicion that the Lonely Planet series has really gone off the boil. I hesitate slightly to rate it bad, because it didn't actively do anything seriously wrong, it was just badly lacking in places. In just a brief trip to Korea staying in just 3 places I noticed some really annoying omissions and contradictions.

For example, the section on the Myeongdong district of Seoul lists very few eating opportunities, giving you the impression that there are no restaurants in this part of town. This is clearly not the case; if you simply walk around enough you'll find plenty of individual streets crowded with perfectly decent restaurants. The guide doesn't need to list all these, it would just be useful if it flagged up that this was the case. I don't need to be spoon fed; I'd much rather be pointed in the right general direction and take it from there, but the Lonely Planet guide doesn't give you that bigger picture. The one listed restaurant I did visit looked dead and uninspiring compared with other places round the corner, so I walked away and found somewhere else.

Similarly, the section on Gyeongju says that there are two trains a day to Seoul (only two?!). What it doesn't bother to mention is that after a shortish train ride to another town along the way you can pick up the KTX high speed train direct to Seoul, giving you the option of a good ten trains a day. Staff at the station will sell you a through ticket. It would have taken just one more sentence to mention this, but the authors didn't seem to think it was worth it (or didn't bother to do their homework). This gives me the impression of a guide that hasn't really been thoroughly field tested.

In the Health section, one paragraph states that no special vaccinations are required or recommended for Korea, but then one page later it states that all travellers to Korea should be vaccinated against Hepatitis A. So, is this a special vaccination? Is it required or not? This kind of ambiguity is really sloppy.

At a general level I find the layout of the books has now got pretty confusing. In any one section key information about a topic such as transport links may be scattered around different subsections, often leaving me wondering "where did I read that?".


There could also be improvements to the indexes in LP guides in general. One of the first things I want to know when I arrive in any country is whether or not it's safe to drink the tap water. This information is generally embedded in the Health section, but why not put an entry in the index pointing directly to this?

Do you know tom? [Login](#) or [Register](#) to add tom to your Network

[Web Feed of tom's Latest Reviews](#)

[RSS WEB FEED](#)

About tom (Tom Heath)



[tom's Home Page](#)

tom's location:
[Borough of Milton Keynes](#)

tom's Interests

[Semantic Web](#)
[ESWC2006 Semantic Web Technologies](#)
[Beer](#)

[RDF Metadata About tom](#)

[RDF META](#)

Figure 12. The author's Revyu profile page, showing review data from the site (left) alongside information from his external FOAF file (right)

In addition, where a user knows another reviewer they can choose to add this person to their social network (as recorded on Revyu). This relationship is then recorded in the triplestore using the `foaf:knows` property. All such triples are exposed in the user's RDF description on the site, allowing them to be combined with other FOAF data from the Web to provide an integrated definition of the user's social network.

6.11.2. Supplementing Film and Book Reviews with External Data

Having determined the `rdf:type` of reviewed books and films using the heuristics described above, and based on the `owl:sameAs` links derived through this process,

additional data is retrieved about the item from external sources and used to supplement reviews with further information about the item.

Where items have been identified as films, information such as the name of the director and the URI of the promotional poster are retrieved by querying the *DBpedia* SPARQL endpoint. This additional information is displayed on the HTML page about the film, alongside reviews of the film that have been created in Revyu, thereby enhancing the value of the site for users without requiring this information to be manually entered into Revyu itself. This mashup of review and film information is illustrated in Figure 13 below.

REYU.COM

ReView ANYTHING

Home | Browse Things | Search Things | Browse People

Login/Register | New Review

The Prestige

Links

Homepage: <http://theprestige.movies.go.com/>
See Also: <http://imdb.com/title/tt0482571/>

Tags

[christian-bale](#) [christopher-nolan](#) [drama](#) [entertainment](#) [film](#) [hugh-jackman](#) [illusion](#) [magic](#) [michael-caine](#) [movie](#) [murder](#) [period](#) [scarlett-johansson](#) [science-fiction](#) [whodunnit](#)

Reviews (1)

★★★★★

by [martinp](#) on 23 Jan 2007


This is a drama about intense rivalry between stage magicians in the late 19th Century. The evocation of the period, although first rate, is not the main attraction, however. The Prestige has an incredibly clever plot including the most ingenious murder I've ever come across. It also has a deeply moving and sad love story hidden in it, which gradually emerges over the course of the film.

The film requires a strong suspension of disbelief on some key points: there is a science-fiction premise which is introduced using the real historical character of Nikola Tesla (I'd rather they had used a fictional scientist). There are a couple more implausibilities required to hold it together (something odd that goes on that none of the characters pick up on and a dead-end that by a huge coincidence turns out not to be a dead-end: I can't be more specific without spoiling the plot).

However, rather than feeling cheated by these aspects of the film, I'm hugely impressed. The writers have taken an implausible (okay, impossible) premise but created an intricate, involving and visual story that would be impossible without that premise. Scenes join up with each other in many subtle ways, echoing the same writers' earlier film Memento. Even when you've seen the twist coming, the final scene which lays it all out are has a lot of impact and I suspect the final shot will haunt my dreams.

I expected the film to be about nice costumes or impressive magical trickery, but it is actually about deep emotions felt by the main characters as they deal with the situations life has dealt them, and it rather than serving up those emotions on a plate, it requires you to think and piece together what you've seen. That's got to be a good thing, in fact the best of what film a be.

The Prestige




directed by [Christopher Nolan](#)

[RDF Metadata About The Prestige](#)

RDF

META

[Write a Review of The Prestige](#)



[Add to del.icio.us](#)

Figure 13. A film review on Revyu (left) shown alongside film data from DBpedia (right)

Similarly links between Revyu and the *RDF Book Mashup* (Bizer, Cyganiak and Gauss, 2007) are exploited as the basis for retrieving book cover and author information which is also then displayed on the Revyu HTML page about the book, as shown in Figure 14.

REYVU.COM

REVIEW ANYTHING

Home | Browse Things | Search Things | Browse People

Login/Register | New Review

The Unwritten Rules of Phd Research, by Gordon Rugg and Marian Petre

Links

Homepage: <http://mcgraw-hill.co.uk/openup/unwrittenrules/>
See Also: <http://www.amazon.co.uk/Unwritten-Rules-Phd-Research/dp/0335213448/>

Tags

[book](#) [guide](#) [phd](#) [research](#) [rules](#)

Reviews (1)


★★★★★

by tom on 07 Nov 2006

Authors Gordon Rugg and Marian Petre tell PhD research as it is, in this essential book for any research student. Whilst there are many books out there about how to tackle research at PhD level, this one feels qualitatively different in the topics it covers, and the honesty with which it does it. My personal favourite section concerns writing style, phrases you may use in your dissertation, and how experienced examiners interpret these phrases. Example; You say: "(though c.f. Green et al (in press) for an interesting re-evaluation of this literature)"; Others read this as meaning: "I've read the advanced literature, so sod off". Brilliant.

What do you think of The Unwritten Rules of Phd Research, by Gordon Rugg and Marian Petre? [Write Your Own Review...](#)

The Unwritten Rules of PhD Research



Gordon Rugg
Marian Petre

The Unwritten Rules of Phd Research by Gordon Rugg and Marian Petre

Open University Press

ISBN: 0335213448

The Unwritten Rules of Phd Research on [Amazon UK](#)

[RDF Metadata About The Unwritten Rules of Phd Research, by Gordon Rugg and Marian Petre](#)




Figure 14. A book review on Revyu (left) shown alongside film data from Amazon/RDF Book Mashup (right)

This approach could be described as using Semantic Web data to produce Web2.0-style mashups at the (human-readable, HTML level, whilst also creating linked data mashups at the RDF level. Not only does this linked data approach to mashups reduce issues with licensing of data for republication, it is also a more Web-like approach; duplicating data is of much lesser value than linking to it, and the user agent of the future should be able to 'look ahead' to linked items and merge data accordingly.

It should be noted that no claims are being made that this form of human-oriented mashup represents something that could not have been achieved using conventional Web2.0 approaches, or provides immediate user benefits over conventional Web2.0 mashups. What distinguishes this approach however is the simultaneous publishing of data and human-oriented mashups, which brings several significant benefits for the developer, for the Semantic Web at large and ultimately for future Web users.

148

Firstly, the development effort is substantially reduced, as a common toolset (e.g. the SPARQL client of the RAP library) can be used to query all data sources, and the ability to easily integrate heterogeneous sources using RDF substantially reduces development costs in producing human-oriented mashups.

Secondly, making and exposing these links in RDF helps to populate the Semantic Web with links between data sets, ensuring that the data integration effort is not lost but can be reused by other parties on the Web.

Lastly, if other sites join Revyu in publishing reviews in RDF, and reference the same URIs, large-scale aggregation of reviews from many sources that would be highly complex using Web2.0 approaches becomes trivial using Semantic Web technologies. The potential then exists to create RDF-based mashups that are infinite in nature, integrating data from arbitrary sources as required and providing a richer and more complete picture to users of how an item has been reviewed across the Web.

6.11.3. Supplementing Reviewed Items by Pre-population

Whilst links from films and books on Revyu to corresponding items in external data sets are created heuristically, a different approach has been followed when linking Revyu to data from the Open Guide to Milton Keynes (Gaved, Heath and Eisenstadt, 2006) and papers from the 6th International Semantic Web Conference (ISWC+ASWC 2007).

The Open Guide to Milton Keynes is a member of the Open Guides family of wiki-based city guides that publish data in RDF. Milton Keynes is a town in south east England, and home of The Open University. Whilst some amenities in the locality, such as pubs and

restaurants, were already reviewed on Revyu, many more were listed in the Open Guide due to its longer history.

Therefore, after identifying items existing in both locations and making the appropriate mappings to avoid duplication, skeleton records were created in Revyu for the remaining items, setting links back to their Open Guide URIs. These skeleton records provide a basic representation of items within Revyu (a title, `rdf:type` statement, keyword tags and links back to the item in the original data source). This serves to encourage users to review items they recognise, ensures greater coverage and consistency of entries than is possible through organic growth, and ensures that items are properly linked across data sources.

These links enable latitude and longitude data for many items to be retrieved from RDF exposed by the Open Guide, and used to show a Google Map of the items location, as shown in Figure 15. The same approach can also be used to expose address, telephone, and opening time information held in the Open Guide, and can be extended to Open Guides for other locations, such as London and Boston.

Revyu.com

Review anything

Home | Browse Things | Search Things | Browse People

Login/Register | New Review

Ye Olde Swan, Woughton On The Green, Milton Keynes


Links

See Also: http://miltonkeynes.openguides.org/?Ye_Olde_Swan%2C_Woughton_On_The_G...

Tags

[bar](#) [beer](#) [food](#) [milton-keynes](#) [pub](#) [restaurant](#) [woughton](#)

Map



Reviews (2)

★★★★☆

by tom on 20 Nov 2006

The Swan (as it's commonly known) is one of the nicer pubs within the Milton Keynes boundary, being situation in the old village (more a large hamlet really) of Woughton on the Green. It's got log fires, a crooked roof, and low ceilings, which all give it a cosy feeling on a winter evening. Come summer time there are plenty of tables outside in a pleasant garden. Given the other options nearby it's a good choice, but has plenty of shortcomings. The pub is heavily geared towards food, which is generally quite good, and this is responsible for a stupidly restrictive policy about where you can sit during busy periods. In practice you may find more than half the pub reserved for diners, making it hard to get a table if you just want a drink. There are a few decent ales on tap, including Deuchars IPA and Old Speckled Hen, which generally seem well kept. Anywhere else I'd give this pub a rating of 3, but in MK the lack of decent choice makes this a 4.

RDF Metadata About
Ye Olde Swan, Woughton On
The Green, Milton Keynes

RDF META

Write a Review of
Ye Olde Swan, Woughton On
The Green, Milton Keynes

Add to del.icio.us

Figure 15. Geodata from the Open Guide to Milton Keynes used to display a Map of a reviewed item's location

This 'pre-population' approach was also used to create skeleton records in Revyu describing papers presented at the 6th International Semantic Web Conference and 2nd Asian Semantic Web Conference (ISWC+ASWC 2007), based on RDF data produced describing the conference (Möller, Heath, Handschuh et al., 2007).

It should be noted that the goal of pre-population from external datasets is not to constrain, but merely to seed users' conceptions of what can be reviewed, where well-defined external data sets exist describing items that may usefully be reviewed in Revyu. Any additional reviews created on the site lead to greater coverage in the trust metrics described in Chapter 7.

151

6.12. *Reusability of Revyu Data*

By making content available in standard formats, Revyu reviews can be syndicated and reused by reviewers who use the site and administrators of third party sites who wish to add value to their existing content by adding review information, or combined with reviews from other sources that are also published in RDF. This can be particularly valuable in overcoming the scenario where an item may not have been reviewed many times on one particular site, but reviews exist elsewhere on the Web.

Multiple routes are provided for accessing and reusing Revyu data. With one line of JavaScript code a user's ten latest reviews can be displayed on a remote Web site. This provides a simple mechanism for syndication of reviews by users who are less technically proficient. More sophisticated syndication options are available via RSS feeds of the latest reviews across the entire site and from each individual user.

Third parties interested in data integration rather than simple syndication have two options: retrieving RDF data from the site by crawling or making one-off HTTP requests; or accessing the data they require via queries to the Revyu SPARQL endpoint.

Revyu exposes data about things, reviews, people, and tags via its SPARQL endpoint, which relies on the RAP SPARQL engine operating against the same MySQL-based triplestore. Providing such a query interface allows third parties to retrieve reviews and related data in a flexible fashion, for reuse in their own applications. Whilst in some ways analogous to Web2.0 APIs which provide remote query capabilities, SPARQL endpoints afford many advantages to the developer: for example, common libraries can be used to query multiple RDF graphs yet return the results as one resultset, effectively allowing joins over multiple data sources.

One example of a third party application that uses the Revyu SPARQL endpoint to retrieve data to enhance its own services is the Semantic Web gateway *Watson* (Aquin, Baldassarre, Gridinoc et al., 2007). *Watson* uses Revyu as a generic reviewing and rating platform, whereby *Watson* users can review ontologies in Revyu and this data is then retrieved via SPARQL queries for display on the *Watson* site. This provides people searching for ontologies with an indication of how particular ontologies are viewed by others, as shown in Figure 16.

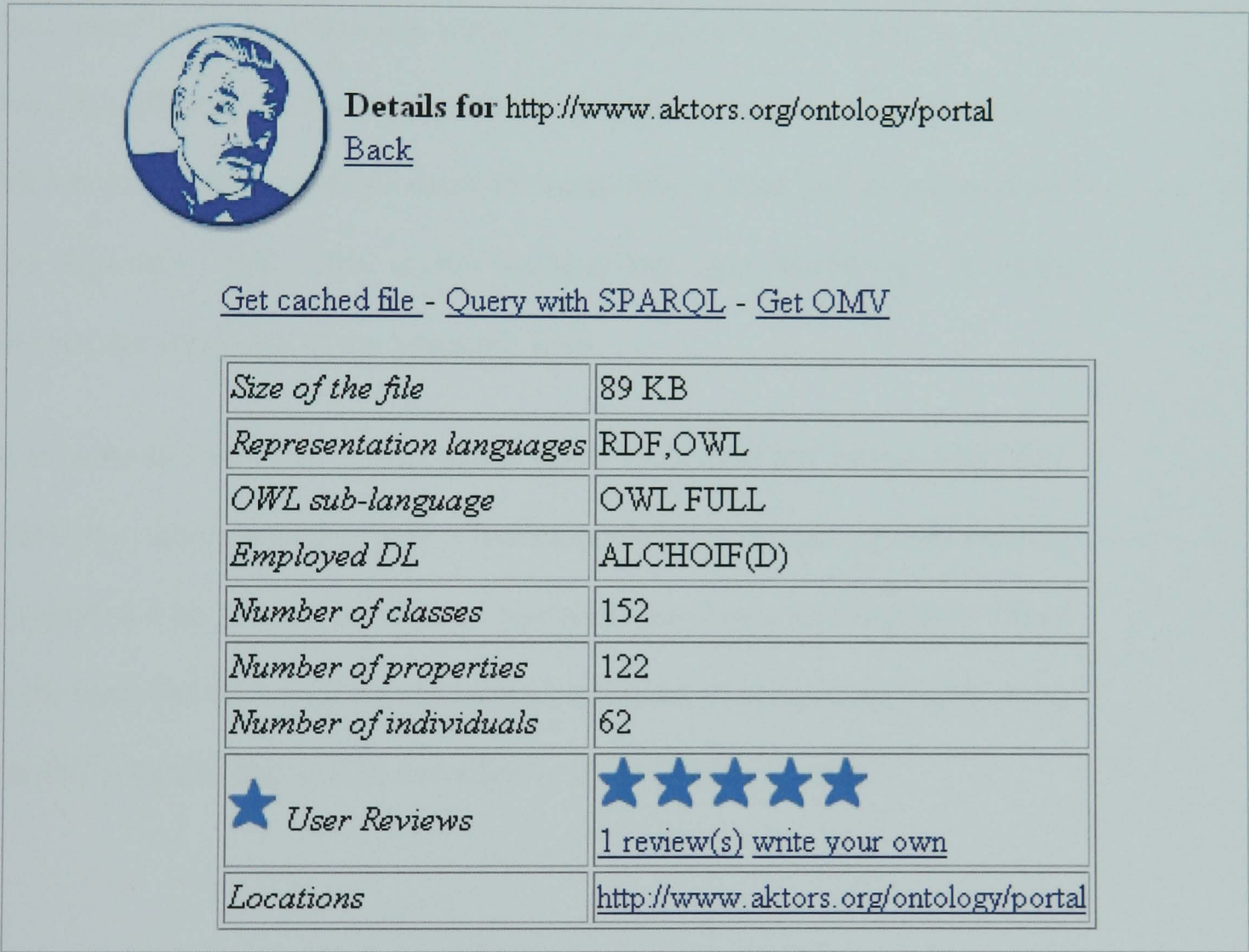


Figure 16. Revyu review data reused in the Semantic Web gateway *Watson*, via the Revyu SPARQL endpoint

6.13. Conclusions

Few mechanisms currently exist that allow non-specialist users to contribute to the Semantic Web. This is in stark contrast to both the conventional Web and Web2.0. Early

growth of the Web is widely attributed to individuals creating personal sites by copying and pasting HTML code. Whilst this approach may not be appropriate to a Semantic Web (novice users may not understand the semantics of statements contained in copied code), Web2.0 applications have demonstrated that regular users can contribute content without specialist skills. With few exceptions, similar tools enabling grassroots publishing on the Semantic Web are not currently available. Revyu is one exception.

Revyu is rare in its status as a publicly available service in daily use that is oriented towards human users but also embodies current best practices in developing for the Semantic Web. By adhering to the well established interaction pattern of completing forms in a Web browser, Revyu allows users to create review data that is immediately usable on the Semantic Web. This occurs without any user knowledge of RDF, ontologies, or even the principles of the Semantic Web.

By providing reviews in a reusable format that is easily integrated and linked with other data, Revyu provides source data that is in a format suitable for computing trust metrics that can be integrated with social networks, as discussed previously in Chapter 5. These metrics form the basis for the social network-based information-seeking approach being investigated in this research, and will be described in the following chapter.

7. Hoonoh: Source-centric Information-seeking with Trusted Social Networks

7.1. Introduction

Chapter 4 reported on a study of how people choose information sources from among members of their social network, and how these choices differ according to characteristics of the task. These insights into the decision-making process in recommendation-seeking form the basis for the two contributions presented in this chapter:

1. Algorithms have been developed that exploit Revyu and other Web data sources to compute trust metrics based on the findings of the study in Chapter 4. These algorithms identify the topics in which individuals have *experience* and *expertise*, and with whom they share *affinity* relationships. Development of these algorithms provides evidence with which to address Research Question 4 ("to what extent can general principles derived from answers to the previous questions be operationalised as computational algorithms that replicate the process of seeking information and recommendations through social networks?").
2. A Web-based system, Hoonoh, has been implemented and deployed that uses these metrics to support source-centric information-seeking within an individual's social network. Hoonoh allows users to search for people with knowledge of particular topics and rank these potential information sources according to the *experience*, *expertise* and *affinity* trust factors. The implementation of Hoonoh provides experience from which answers to Research Question 5 ("How feasible

is the implementation of user-oriented systems that exploit such algorithms?") can be derived. The system is presented in this chapter, whilst the answer to this question is discussed in Section 9.3.2.

The first half of this chapter details the Hoonoh algorithms for computing *experience*, *expertise* and *affinity* metrics, also describing the data used as input to the algorithms, the technical infrastructure on which they are implemented, and the ontology with which the resulting metrics are described. The second half presents Hoonoh from a user perspective, and describes the technical implementation of the system.

An evaluation of the results that Hoonoh provides based on the underlying Hoonoh algorithms is reported in Chapter 8, in order to address Research Question 6 ("If such systems can be implemented, how do they perform relative to human performance of equivalent tasks?")

7.2. Computing Knowledge and Trust Relationships: The Hoonoh Algorithms

The set of Hoonoh algorithms consists of algorithms for generating 'experience', 'expertise' and 'affinity' metrics. These metrics represent respectively the predicted trustworthiness of an individual with regards to a topic, based on his or her experience of and expertise in that topic (person \rightarrow topic relationships), and the predicted trustworthiness of an individual based on the affinity relationship between the information seeker and that individual (person \rightarrow person relationships). These factors were chosen for the reasons outlined in Section 5.2.1.

A fundamental aspect of my work, based on the findings reported in Chapter 4, is the principle that trust can be topical; one person may be highly trusted for recommendations in one domain but trusted very little in others. For example, one may trust a friend who is a banker to give sound financial advice, but never trust her film recommendations. This trust topicality is supported by the experience and expertise algorithms, whilst affinity captures a more universal trust relationship from one individual to another that is not topical in nature (see Section 4.6.4 for more details).

7.2.1. Input Data to Hoonoh Algorithms

As discussed in Section 5.2.2, the Hoonoh algorithms seek to exploit distributed data sources in order to compute experience, expertise and affinity metrics. Rating data from reviews entered into Revyu is the primary data source used by the algorithms, as the judgements embodied in such ratings provide a basis for computing affinity metrics between individuals and metrics for expertise, as will be discussed below. Revyu-based experience measures are supplemented by data from third-party Web2.0 and Semantic Web sources. Details of how this is achieved are given in Section 7.2.5.

7.2.2. Tags as Topics

Keyword tags used in Revyu seed the list of topics in which individuals may have experience or expertise, and also provide a basis for computing measures of experience. Each keyword tag is taken to denote one topic; lexical variations in keywords, synonyms and homonyms are ignored for the purposes of this research. This approach represents the best available compromise of usability and comprehensiveness of topical coverage: requiring non-specialist users to navigate large ontologies in order to 'semantically tag' reviewed items is likely to present a significant usability barrier. Section 9.3.4 discusses

this compromise in a little more detail, and highlights ways in which this aspect may be developed in future work.

7.2.3. Proxy Metrics

The Hoonoh algorithms are directly informed by the study in Chapter 4. However, developing algorithms that fully represent the trust factors has not been possible in the cases of experience and expertise. Instead, measures that serve as proxies for these trust factors have been developed.

For example, computing a true *expertise* score in any one domain is problematic. Expertise was defined in Chapter 4 as "*the source has relevant expertise of the domain of the recommendation-seeking; this may be formally validated through qualifications or acquired over time.*"

Appropriate sources of background knowledge indicating qualification in a domain are not readily available on the Web. Where they are available they tend to be widely distributed according to the domain of the qualification, and are not generally available in structured, machine-readable form. For example, one's family doctor may be qualified in general medical practice; however, evidence of this in the form of a machine-readable certificate of qualification and competence from a recognised medical authority is not available on the Web.

Consequently, I have developed a metric called *credibility* that serves as a proxy for expertise. This credibility algorithm emphasises the more socially-constructed and endorsement-oriented aspects of expertise. Whilst formal qualifications can serve as a significant indicator of domain expertise, in many cases the status of 'expert' will be a

product of domain knowledge and the endorsement of this knowledge by a wider community.

The algorithm is detailed below, however in summary: an individual is deemed credible with respect to a particular topic if their ratings of items related to that topic are validated and endorsed by the community as a whole, through strong correlations with other ratings of that item. Therefore, while it reflects a pragmatic decision based on available data, it is argued that the credibility algorithm captures a substantial proportion of the notion of expertise. The efficacy of credibility as a proxy for expertise is discussed in more detail in Sections 8.6 and 9.3.3.

Asking users to rate the utility or value of specific reviews³⁵ was rejected as a means to establish the credibility or domain expertise of individuals. As Dieberger, Dourish et al. (2000) observe, "an 'expert reviewer' is not the same as a 'domain expert'." (pp. 42) In order to confidently integrate such data in metrics it would be necessary to establish whether such meta-reviews serve simply as a measure of some characteristics of the review itself (perhaps reflecting the quality of the writing or level of detail provided), rather than as a reliable measure of the reviewer's expertise in the domain of the reviewed item.

In the case of experience, large volumes of data are available on the Web that may indicate an individual's experience of a particular domain. For example, where someone has rated a large number of hotels on a travel Web site, it may be concluded that they have substantial experience of the 'hotel' domain. However, automatically validating with

³⁵ In the style of Amazon's "Was this review helpful?" meta reviews.

a high degree of confidence that this is the case is non-trivial. Therefore, I have developed a more conservative metric called 'usage', detailed below, that primarily reflects an individual's frequency of use of keyword tags. This serves as a proxy measure for experience of particular domains.

The confidence with which experience can be inferred from tag usage varies across sources of tagging data. For example, an individual may have bookmarked a large number of Web sites at *Del.icio.us* using particular keyword tags. This may indicate that he or she has some experience of the topics denoted by the tags, or simply reflect the gathering of relevant material in order to research a new topic of which he or she currently has no experience.

Usage metrics based on tagging data from Revyu can be considered more reliable than those from *Del.icio.us*, as tags can only be used in conjunction with submission of a review. This increases the likelihood that an individual does in fact have experience of the topics represented by the tags, as submission of a review can be assumed to be predicated on some experience of the item being reviewed. To elaborate on the previous example, in the course of researching potential holiday destinations a user may bookmark many resources using the tag 'hawaii', but eventually choose to visit Mexico instead. In contrast, where a user has reviewed an item there is a reasonable likelihood that they have some experience of the topics denoted by that item's tags.

No proxy metric is required to represent affinity, as comparing ratings between individuals allows for computation of affinity metrics with a reasonable degree of confidence. Because Revyu accepts reviews of any type of item, the data on which

affinity is calculated reflects value judgements from a range of domains, thereby ensuring that affinity metrics capture more than simply taste overlap.

7.2.4. Computing Trust Metrics from Revyu Data

The following sections detail each of the Hoonoh algorithms. Summary statistics are shown in Table 3, Table 4 and Table 5 to illustrate the distribution of values in the output of each algorithm. These statistics were generated from a snapshot of the Revyu data set taken on 11.11.2007. This set was subjected to a small amount of data cleaning to remove duplicate reviews and typing errors in tags, and to ensure consistency in the syntactic form of tags used by reviewers (e.g. replacing 'miltonkeynes' with 'milton-keynes'). After cleaning the data set consisted of 571 reviews written by 139 reviewers, from which the algorithms produced a final Hoonoh data set of 25509 triples.

Usage Algorithm (Experience)

Broadly speaking this algorithm generates person \rightarrow topic usage metrics by calculating what proportion of all items tagged with a particular tag that person has reviewed. The algorithm is shown as pseudo-code in Figure 17.

```

$tags = all tags in Revyu
for each ($tags as $tag) {
    $things = all things tagged with $tag
    $reviewers = all reviewers who have reviewed at least one of $things
    for each ($reviewers as $reviewer) {
        $count[$reviewer][$tag] =
            count(number of $things reviewed by $reviewer)
    }
    $maxcount = highest($count)
    for each ($count[$reviewer][$tag]) {
        $usage[$reviewer][$tag] =
            $count[$reviewer][$tag] / ($maxcount + $constant)
    }
}

```

Figure 17. Usage (Experience) algorithm in pseudo-code

The $\$usage[\$reviewer][\$tag]$ score provides a relative measure of an individual's experience of a topic, based on data available within Revyu, and can be represented by the following equation, where e stands for the usage score, c for the reviewer's tag count (i.e. $\$count[\$reviewer][\$tag]$), m for the $\$maxcount$ and k for the constant, then:

$$e \leq \frac{c}{m+k}$$

Consequently, the usage score e will always have a value greater than 0 and less than 1. Reviewers who have not reviewed items tagged with a particular tag are simply excluded from the calculations for that tag, making values of 0 impossible. Furthermore, assuming

that the constant in the algorithm is greater than 0, then even where the value of c is equal to the value of m then the value of a usage score must be less than 1.

In general, the role of the constant k in the equation above is to mediate the effects on usage scores of low numbers of reviews related to a particular topic. For example, where a tag exists that has only been associated with one item, which has in turn been reviewed by only one person, the maximum usage score the reviewer can receive for that topic is defined by:

$$e_{\max} = \frac{1}{1+k}$$

By default the value of k is set to 4, giving a maximum possible usage score of 0.2 in this scenario. As the number of reviews of items related to a particular topic increases (and the accuracy of usage metrics presumably increases due to the larger amount of data on which to base the computation) the constant k has relatively less impact.

One other feature of the algorithm worth noting is that people can get credit for experience of topics for which they have never used the corresponding tag. For example, if one person reviews a hotel and simply tags it 'hotel' but a second person tags it 'hotel' and 'accommodation', the first will also receive a usage score for the topic 'accommodation' as both reviews are of the same item. This helps ensure a broader spread of experience scores across related topics and mitigates possible negative effects of different individuals using different terminology when referring to the same topic.

Usage Score	Frequency	Usage Score	Frequency	Usage Score	Frequency
0.9420	1	0.2857	6	0.0769	1
0.9070	1	0.2727	2	0.0741	10
0.8889	1	0.2500	6	0.0714	6
0.8824	1	0.2222	5	0.0698	3
0.8710	1	0.2143	1	0.0667	14
0.8667	1	0.2083	2	0.0588	2
0.8519	1	0.2000	1967	0.0580	1
0.8400	1	0.1852	2	0.0556	10
0.8333	2	0.1818	4	0.0500	8
0.8000	1	0.1667	224	0.0465	8
0.7778	1	0.1481	1	0.0435	3
0.7333	2	0.1429	50	0.0417	30
0.7143	2	0.1333	3	0.0400	6
0.6923	1	0.1250	30	0.0370	6
0.6667	5	0.1111	24	0.0333	12
0.6364	6	0.1000	8	0.0323	6
0.6000	5	0.0930	1	0.0294	11
0.5556	6	0.0909	16	0.0290	10
0.5000	17	0.0882	1	0.0278	12
0.4286	37	0.0833	23	0.0233	25
0.3333	171	0.0800	2	0.0145	18

Table 3. Distribution of usage (experience) scores in a sample Hoonoh data set

The large number of experience scores of 0.2000 primarily reflects the number of cases where someone has reviewed just one item related to a particular topic, giving a usage score e of:

$$e = \frac{1}{1 + 4} = 0.2$$

Credibility Algorithm (Expertise)

The credibility algorithm computes person → topic credibility metrics by comparing the numerical rating component of each review to the mean rating of that item across all users. A mean is then taken of all a reviewer's review-specific credibility scores for items tagged with a particular tag, to produce a reviewer's credibility score for that topic. A pseudo-code representation of the algorithm is shown in Figure 18.

```
$tags = all tags in Revyu
for each ($tags as $tag) {
    $things = all things tagged with $tag
    for each ($things as $thing) {
        $reviews = all reviews of $thing
        if (count($reviews) > 1) {
            $meanrating = mean($reviews['rating'])
            for each ($reviews as $review) {
                $ratingdistance[$review] =
                    absolutevalueof($meanrating-$review[rating])
                $adjustedratingdist[$review] =
                    $ratingdistance[$review]/($ratingmax-1)
                $credibility[$review] = 1-$adjustedratingdist
                $credibilitysum[$reviewer] += $credibility[$review]
                $count[$reviewer]++
            }
        }
    }
    $credibility[$reviewer][$tag] =
        $credibilitysum[$reviewer]/$count[$reviewer]
}
```

Figure 18. Credibility (Expertise) algorithm in pseudo-code

In this implementation of the algorithms, the variable `$ratingmax` is in fact a constant with a value of 5, representing the maximum possible rating that can be given to an item in Revyu. This constant is an essential component of the algorithm as it allows the 'rating distance' of a particular review to be adjusted to a value of less than or equal to 1, which simplifies the combination of trust metrics across multiple factors. Whilst the maximum possible value for credibility is 1, the lower bound for credibility values tends to zero

(but cannot reach 0), as the rating distance for a particular item can never mathematically be equal to the value of $\$rating_{max}-1$.

One additional constraint hardwired into the algorithm is the exclusion of items that have only been reviewed once. This prevents the assignment of credibility values of 1 to individuals whose reviews have not been validated by any other reviewers.

Credibility Score	Frequency	Credibility Score	Frequency	Credibility Score	Frequency
1.0000	625	0.8906	1	0.7918	1
0.9792	1	0.8834	1	0.7834	1
0.9688	2	0.8792	1	0.7813	1
0.9500	1	0.8750	159	0.7500	149
0.9375	18	0.8542	2	0.7292	1
0.9197	1	0.8500	14	0.6875	1
0.9168	57	0.8438	2	0.6668	51
0.9167	4	0.8417	1	0.6250	31
0.9028	1	0.8334	1	0.5833	5
0.9000	26	0.8333	46	0.5000	30
0.8958	3	0.8250	1	0.3333	17
0.8947	1	0.8125	5		

Table 4. Distribution of credibility (expertise) scores in a sample Hoonoh data set

The high number of credibility scores of 1 shown in Table 4 reflects the relatively low density of the sample data set. The 571 reviews referred to 466 unique things, giving a mean number of ratings per item of 1.23. As a result, where one thing was reviewed by two people who gave the item the same rating, each would get a credibility score of 1 for each topic associated with the reviewed item, unless they have provided ratings of other things with the same topic and with which other people disagree.

Unlike the usage algorithm, the credibility algorithm does not produce metrics that are mediated by the overall number of things that have been tagged with a particular tag, or the number of reviews of these things. In the case of usage these mechanisms serve an important function in ensuring more balanced usage scores. However, in the case of credibility, metrics are by definition mediated by the ratings of others and therefore must

reflect the conclusions that can be drawn about a user's credibility based on the available data.

In common with usage, the credibility algorithm enables metrics to be generated for people regarding topics for which they have never used the corresponding tag, as long as they have reviewed an item which another user have tagged with the appropriate tag.

Affinity Algorithm

Whether or not an affinity exists between two individuals is determined by a combination of the following factors derived from the reviews they have submitted to Revyu: the extent to which both parties have rated the same items (i.e. the overlap in rated objects) which is referred to here as the 'item overlap'; and the consistency in the ratings given by each party to items both have reviewed; this is referred to as the 'rating overlap'. The affinity algorithm is presented in pseudo-code in Figure 19.

```

$reviewers = all reviewers in Revyu
for each ($reviewers as $reviewer) {
    $others = all $reviewers excluding $reviewer
    $highestitemoverlap = 0 // for highest item overlap between all users
    for each ($others as $other) {
        $overlappingitems = all items reviewed by both $reviewer and $other
        if (count($overlappingitems) > 0) {
            if (count($overlappingitems) > $highestitemoverlap) {
                $highestitemoverlap = count($overlappingitems)
            }
            for each ($overlappingitems as $item) {
                $ratingdistance[$item] =
                    absolutevalueof($reviewer[rating]-$other[rating])
                $ratingdistancesum[$reviewer][$other] += $ratingdistance[$item]
            }
            $meanratingdistance[$reviewer][$other] =
                $ratingdistancesum[$reviewer][$other]/count($overlappingitems)
            $adjustedratingdist[$reviewer][$other] =
                $meanratingdistance[$reviewer][$other]/($ratingmax-1)
            $ratingoverlap[$reviewer][$other] =
                1-$adjustedratingdist[$reviewer][$other]
            $affinity[$reviewer][$other] =
                $ratingoverlap[$reviewer][$other] *
                (count($overlappingitems)/$highestitemoverlap)
        }
    }
}

```

Figure 19. Affinity algorithm in pseudo-code

As with credibility, the variable `$ratingmax` is a constant of value 5, representing the maximum possible rating that can be given to an item in Revyu.

In the algorithm, rating overlap is mediated by the ratio of item overlap between two users to the highest item overlap between any users in the system (`$highestitemoverlap`). This adjustment avoids false positives in affinity scores resulting from low item overlaps. For example, two reviewers may have reviewed one item in common (`count($overlappingitems) === 1`) and given the same or very similar ratings; this would result in `$ratingdistance[$item]` being low, and overall `$ratingoverlap[$reviewer][$other]` being high. Whilst the two reviewers may happen to agree in their ratings of this item, in reality they may have a low overall affinity, which would not become apparent without reviewing more items in common. Taking relative item overlap into account mitigates this effect. The overall outcome of this may be described as: reduce every affinity score by an amount that is inversely related to the number of overlapping items on which the affinity score is based.

Affinity metrics can be in the range 0-1 inclusive. Values of 0 result from situations where two reviewers disagree to the greatest possible extent in their ratings of all overlapping items (i.e. one gives a rating of 1 while the other gives a rating of 5). This produces a mean rating distance (`$meanratingdistance`) of 4, an adjusted rating distance (`$adjustedratingdist`) of 1, and a rating overlap (`$ratingoverlap[$reviewer][$other]`) of 0. Therefore, irrespective of the number of overlapping items the affinity will always be 0. An affinity score of 1 is only possible where two reviewers agree fully in all their reviews of overlapping items, and also have the `$highestitemoverlap` in the system.

It should also be noted that affinity relationships are symmetric, and that metrics are computed for all pairs of reviewers in the system who share an item overlap, irrespective of whether or not they know each other.

Affinity Score	Frequency	Affinity Score	Frequency	Affinity Score	Frequency
0.8500	2	0.5000	2	0.1500	56
0.7500	2	0.4000	10	0.1000	32
0.6000	2	0.3500	6	0.0500	14
0.5500	4	0.2000	90	0.0000	6

Table 5. Distribution of affinity scores in a sample Hoonoh data set

The high number of scores of 0.2000 reflects the fact that there are a high number of affinity relationships (90) between reviewers which are based on an `itemoverlap` of 1 and complete agreement in ratings (i.e. a `ratingoverlap` of 1). The highest `itemoverlap` in the sample dataset is 5, giving 90 affinity scores of 0.2000.

7.2.5. Supplementing Experience Metrics through Additional Data Sources

In order to increase the range of topics for which users in the system have experience scores, and the number of users represented in the system, the usage metric generation process takes into account data from sources other than Revyu, such as *Del.icio.us* tagging data and background Semantic Web data. In both cases these data sources are used to create `person → topic` usage scores where they do not already exist, or raise existing usage scores to a minimum level.

Usage Scores from Del.icio.us Tagging Data

From a FOAF-oriented crawl of Semantic Web data totalling 6 million RDF triples, a number of individuals were identified who have *Del.icio.us* accounts and choose to publish their `foaf:mbox_sha1sum`. Using the account username as published in their

FOAF data, the tags these individuals have used to bookmark Web sites are retrieved from *Del.icio.us* in 'JSON' format.

For each tag that has been used a minimum number of times (the usage threshold is currently set relatively arbitrarily at 10) the user is assigned a standard nominal experience score (currently 0.1) for that topic. This value is constant irrespective of the frequency of usage of the tag above the threshold, in recognition that tag usage may not be strongly correlated with real experience of the topic (as discussed above) and therefore caution is required.

Where a user has an existing usage score for a particular topic that exceeds the nominal score assigned based on *Del.icio.us* tag usage, the existing score stands unchanged. Where they have an existing score that is lower than the nominal score, this is increased to equal the nominal score. No attempt is made to supplement Revyu-derived credibility and affinity metrics based on *Del.icio.us* data, as bookmarks do not carry ratings, endorsements, or other value judgments from which these may be derived.

Unfortunately, tag usage data from all users of *Del.icio.us* can not be used to generate usage metrics, as *Del.icio.us* does not associate its users with globally unique identifiers (such as `foaf:mbox_sha1sum`) that are necessary for integration with social networks in Hoonoh.

Where relationships can be found between an individual's `foaf:mbox_sha1sum` and their username on other tagging services, such as *Flickr*, there is the potential to extend this process in order to further increase the coverage of usage metrics in Hoonoh. Where individuals have shared their photos online through *Flickr* and have tagged these with place names or locations, this could be used to infer that they have some experience of

this location. This may represent a more robust indication of experience than tags associated with *Del.icio.us* bookmarks, and may be investigated in future work. Unfortunately, whilst *Flickr* does provide an API through which tagging data can be accessed, creating numerous wrappers for proprietary APIs is less scalable as a data gathering mechanism compared to using Semantic Web data which can be accessed using standard technologies.

Usage Scores from background Semantic Web Data

Data sets are beginning to emerge on the Semantic Web which can be mined to identify experience relationships between individuals and topics. For example, conferences in the Semantic Web field regularly publish data about the event itself in RDF (Möller, Heath et al., 2007). In the case of the 3rd European Semantic Web Conference (ESWC2006), this included a 'semantic delegates list' published in RDF, in which those attending the conference could choose to be included (Heath, Domingue et al., 2006).

As a proof of concept, this list has been used to generate additional experience relationships within Hoonoh (or supplement those that already exist), linking individuals to topics of which they are likely to have some experience having attended the conference. These topics are: *conference*, *budva* and *montenegro* (the town and country where the conference was held), *semantic web* (the theme of the conference), *eswc2006* (the abbreviation often used to refer to the event).

7.3. *Technical Implementation of the Hoonoh Algorithms*

The Hoonoh algorithms are implemented within the Trust Computation Subsystem, previously shown in Figure 4.

Execution of the Hoonoh algorithms requires the processing of potentially large amounts of data, primarily from Revyu but potentially from many sources across the Web. The initial technical approach used for implementing the Hoonoh algorithms was to make a relatively high number of queries to the Revyu SPARQL endpoint to retrieve relevant data in relatively small amounts; this data was then processed in memory by a number of PHP scripts to compute experience, expertise and affinity trust metrics which were then stored as persistent RDF triples in the Hoonoh triplestore.

Testing this approach on an earlier Revyu data set of a few hundred reviews revealed that the approach did not perform adequately even on a data set of this size, and consequently would not scale to the larger number of reviews that has since been accumulated. The primary performance bottlenecks were the response speed of the Revyu SPARQL endpoint (since improved due to developments in the underlying RAP library), the large number of queries made to the endpoint, and the large number of array manipulations required in the PHP scripts. Many of these array manipulations were made necessary by the lack of aggregate functions such as *COUNT* in the SPARQL query language.

To address these limitations a decision was made to introduce a caching layer between the data sources and the live Hoonoh triplestore. This caching layer provides a higher performance data source against which the Hoonoh algorithms can be executed, and can be populated by a small number of SPARQL queries to Revyu and HTTP requests to

data sources such as *Del.icio.us*. This reduces the load on data sources and allows the process of generating trust metrics to be carried out offline, with the results merged into the live Hoonoh triplestore on completion. The source data cache itself, the cache population scripts, the scripts that implement the Hoonoh algorithms, and those that populate the live Hoonoh triplestore make up the Trust Computation Subsystem shown in Figure 5.

The platform chosen for the source data cache was the MySQL database server although any enterprise-class relational database management system would be suitable. Somewhat against the Semantic Web-oriented approach of this research, the cache population scripts retrieve data from the various data sources and manipulate it for storage as relational data in a conventional, normalised MySQL database. This allowed the Hoonoh algorithms to be implemented primarily as a series of SQL queries, coordinated by one primary PHP script that oversees the caching of data and execution of the algorithms.

Whilst SQL may not appear an obvious choice for implementing algorithms such as these, it does provide a highly optimised environment for queries that involve many joins, and has the benefit of many aggregate and mathematical functions (such as *COUNT*, *AVG* and *ABS*, that respectively allow for counting of results, calculating the mean of results, and returning the absolute value of a result) which are absent from SPARQL and would otherwise require computation at the (less optimised) PHP level. Future additions to the SPARQL query language and wider availability of high performance triplestores may allow an RDF-based caching layer to replace the current relational approach.

The current approach to generating Hoonoh metrics, based on a relational cache and implementation of the algorithms in SQL, has reduced the total time required for all Revyu-dependent operations to less than 40 seconds, compared to over one hour with the previous approach. The only remaining bottleneck is the insertion of generated triples into the live Hoonoh triplestore; however this reflects a shortcoming of the method RAP uses to read triples into database-backed stores and could be resolved by migration to a different platform.

7.4. *Representing Computed Trust Relationships*

Once computed, trust relationships based on these metrics are stored in the Hoonoh triplestore, according to the Hoonoh ontology³⁶. The ontology models person \rightarrow topic and person \rightarrow person relationships based on all five trust factors identified in the empirical study presented in Chapter 4.

Nine classes are defined in total. Five of these are used to express trust relationships in the Hoonoh triplestore.

ExperienceRelationship, ExpertiseRelationship, and ImpartialityRelationship represent person \rightarrow topic relationships (as such they are subclasses of the TopicalRelationship class), whilst AffinityRelationship and TrackRecordRelationship represent person \rightarrow person relationships (these are subclasses of the InterpersonalRelationship class).

³⁶ <http://hoonoh.com/ontology>

`TopicalRelationship` and `InterpersonalRelationship` are not intended to be used to describe instance data but are provided simply as unifying superclasses, and are themselves subclasses of a unifying `Relationship` class. A class for people is not defined in the Hoonoh ontology as the `Person` class from the FOAF ontology (`foaf:Person`) is reused.

Trust relationships in Hoonoh are modelled as instances of classes, in order to allow varying degrees of trust to be quantified by specifying numerical values as properties of the relationships. This is achieved using the `hoonoh:value` property, which has an `rdfs:domain` of `hoonoh:Relationship` and an `rdfs:range` of `xsd:decimal`³⁷. This modelling pattern was deemed preferable to modelling trust relationships as binary, given that trust relationships are being computed based on numerical data. Inferring binary relations from such data would still require the setting of an arbitrary numerical threshold at which to set a relationship; therefore it was deemed preferable to expose numerical values for trust relationships and allow applications to interpret these as desired.

The person from whom the relationship originates (the 'source') is identified using the `hoonoh:from` property, which has a domain of `hoonoh:Relationship` and a range of `foaf:Person`.

The topic to which experience, expertise and impartiality relationships relate is defined using the `hoonoh:toTopic` property, which has a domain of `hoonoh:Relationship` and

³⁷ The `hoonoh:` prefix refers to the base URI of the Hoonoh ontology ('<http://hoonoh.com/ontology#>'), the `rdfs:` prefix to the RDF Vocabulary Description Language (RDF Schema), and the `xsd:` prefix to XML Schema.

a range of `hoonoh:Topic`, itself a subclass of the `Concept` class from the SKOS Vocabulary (Miles and Brickley, 2005). Conversely, the description of affinity and track record relationships is completed by use of the `hoonoh:toPerson` property which defines the individual who to whom the relationship refers. This property has a domain of `hoonoh:Relationship` and a range of `foaf:Person`.

Figure 20 provides a schematic view of how an `ExpertiseRelationship` is modelled in the Hoonoh ontology.

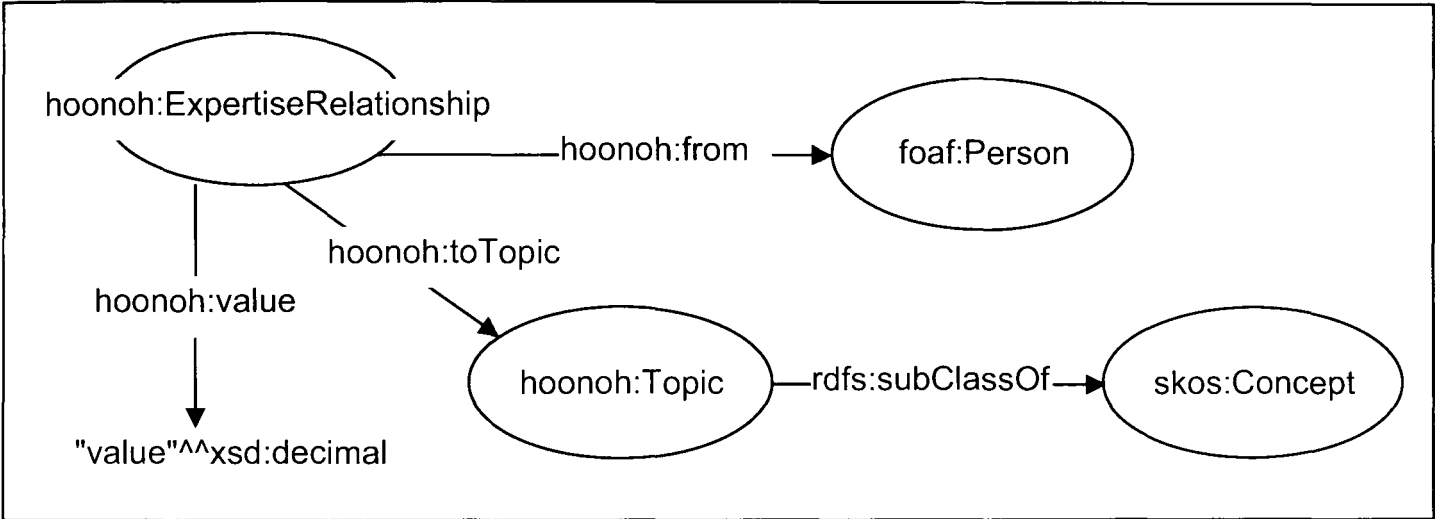


Figure 20. Schematic diagram showing the relationships between classes and properties in the Hoonoh ontology

To complement the schematic view, Code Fragment 5 and Code Fragment 6 show examples of how an `ExpertiseRelationship` and an `AffinityRelationship` can be modelled using the Hoonoh ontology³⁸.

³⁸ Note that these examples are fictional and have been deliberately constructed with URIs that are shorter than those used in the live Hoonoh site, in order to improve readability of the code fragments.

```
<?xml version="1.0" encoding="UTF-8" ?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:hoonoh="http://hoonoh.com/ontology#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xml:base="http://hoonoh.com/">

  <hoonoh:ExpertiseRelationship
    rdf:about="relationships/expertise/abc123/example">
    <hoonoh:from rdf:resource="people/abc123"/>
    <hoonoh:toTopic rdf:resource="topics/example"/>
    <hoonoh:value
      rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.7292
    </hoonoh:value>
  </hoonoh:ExpertiseRelationship>

  <foaf:Person rdf:about="people/abc123">
    <foaf:mbox_sha1sum>abc123</foaf:mbox_sha1sum>
  </foaf:Person>

  <hoonoh:Topic rdf:about="topics/example">
    <rdfs:label>example</rdfs:label>
  </hoonoh:Topic>

</rdf:RDF>
```

Code Fragment 5. An example Expertise relationship described using the Hoonoh ontology

```

<?xml version="1.0" encoding="UTF-8" ?>
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
  xmlns:xsd="http://www.w3.org/2001/XMLSchema#"
  xmlns:owl="http://www.w3.org/2002/07/owl#"
  xmlns:hoonoh="http://hoonoh.com/ontology#"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xml:base="http://hoonoh.com/">

  <hoonoh:AffinityRelationship
    rdf:about="relationships/affinity/abc123/xyz789">
    <hoonoh:from rdf:resource="people/abc123"/>
    <hoonoh:toPerson rdf:resource="people/xyz789"/>
    <hoonoh:value
      rdf:datatype="http://www.w3.org/2001/XMLSchema#decimal">0.8500
    </hoonoh:value>
  </hoonoh:AffinityRelationship>

  <foaf:Person rdf:about="people/abc123">
    <foaf:mbox_sha1sum>abc123</foaf:mbox_sha1sum>
  </foaf:Person>

  <foaf:Person rdf:about="people/xyz789">
    <foaf:mbox_sha1sum>xyz789</foaf:mbox_sha1sum>
  </foaf:Person>

</rdf:RDF>

```

Code Fragment 6. An example Affinity relationship described using the Hoonoh ontology

The Hoonoh triplestore hosts the trust relationship data for the Hoonoh system described below. This data is also republished on the Web as crawlable RDF and via a SPARQL endpoint³⁹ for potential reuse in other applications. It is worth noting that the Hoonoh triplestore is oriented specifically towards storing trust relationship data, not generic information (such as a names or home page addresses) about individuals who have trust relationships generated by the Hoonoh algorithms. Instead this information is harvested from the Web and stored in a dedicated triplestore, as described in Section 7.5.6 below.

³⁹ <http://hoonoh.com/sparql>

7.5. Hoonoh

Hoonoh is available online at <http://hoonoh.com/>. The aim of the system is to help the user identify individuals who may have knowledge about a particular topic or topics, from among members of his or her social network. Figure 21 shows the Hoonoh homepage.

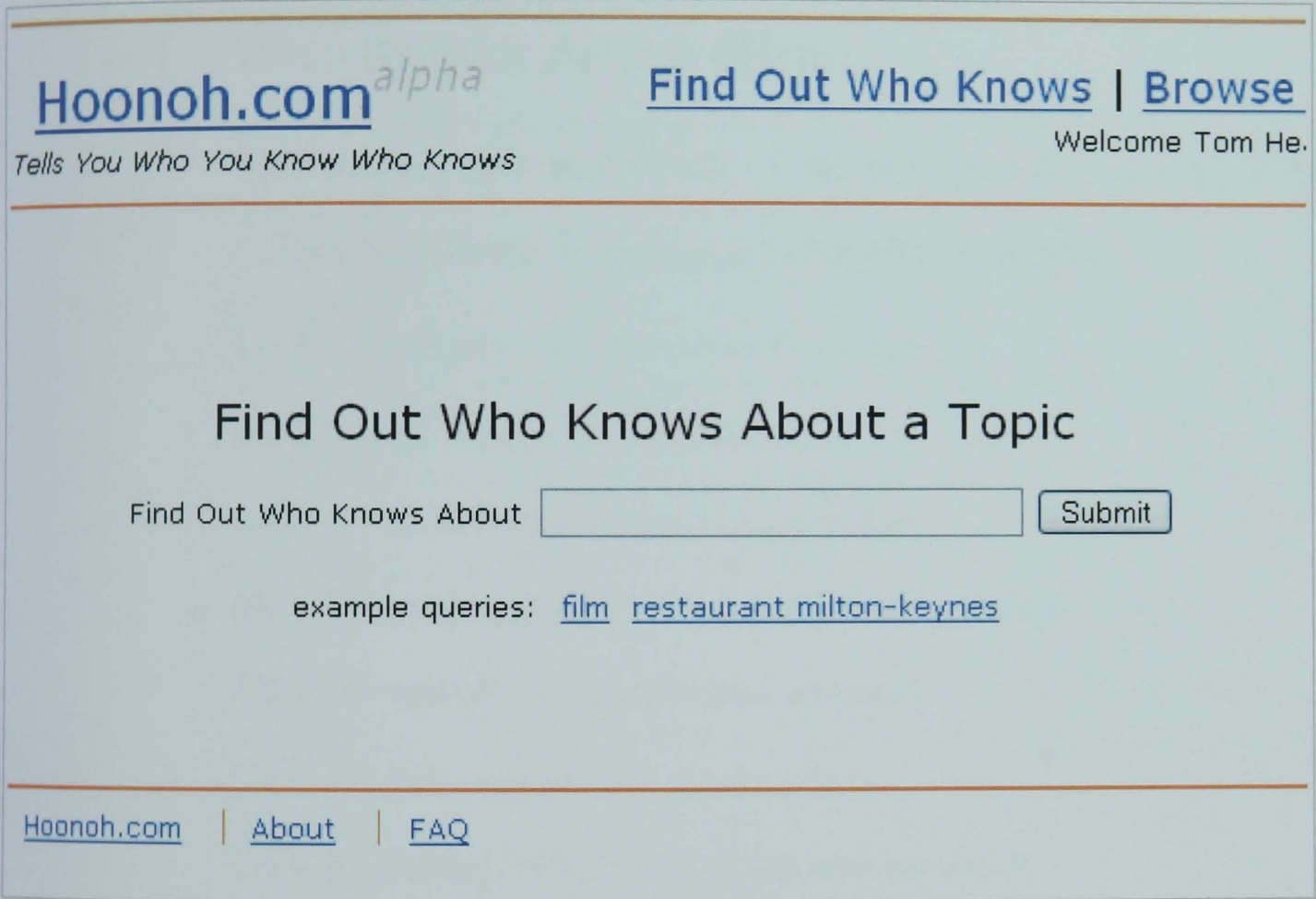


Figure 21. The Hoonoh Home Page

7.5.1. Hoonoh from the User Perspective

The system is designed to function in much the same way as a conventional search engine, where the user specifies the topic of the query in the form of keywords. However, rather than returning a list of ranked documents, Hoonoh returns a list of people from the user's social network who have some knowledge of the topics specified by the query. Where more than one source is identified, Hoonoh enables the user to rank the

individuals according to the experience, expertise and affinity trust factors, as shown in Figure 22 and Figure 23. The details of how these factors are employed are described in Section 7.5.2 below.



Figure 22. Example output: the author's social network ranked by expertise in the topic 'film'

Comparing Figure 22 and Figure 23 it can be seen that the top result when the network is ranked by expertise ('Danni') is ranked last when affinity is taken into account.

Who Knows About **film**?

Limit Results to: Friends (55) + Friends of Friends + Friends of Friends of F

Weight Results by: [Experience](#) | [Expertise](#) | [Affinity](#) |

- [1] 0.683 **Crash** - what do they know about film?
-
- [2] 0.675 **drewp** - what do they know about film?
-
- [3] 0.592 **hockeyshooter** - what do they know about film?
-
- [4] 0.242 **Fin** - what do they know about film?
-
- [5] 0.142 **cancer** - what do they know about film?
-
- [6] 0.042 **Enrico Motta** - what do they know about film?
-
- [7] 0.042 **Danni** - what do they know about film?

Figure 23. Example output: the author's social network ranked by a combination of experience in the topic 'film' and affinity

These 'people results' are then complemented by details of items these people have reviewed on Revyu (Figure 24), allowing the user to see both the trusted sources within their network, and items that may provide a solution to their query.

What Does *drewp* Know About *film*?

film Reviews

Review of: Idiocracy

5/5 on 29 March 2007

Very silly, lots of great satire. ([More of Idiocracy... at Revyu.com](#))

Review of: You, Me and Dupree

4/5 on 24 November 2006

Pleasant formula comedy. I imagine I'd watch it again (if I ever watched movies c the theater) ([More of You, Me and Dupr... at Revyu.com](#))

Review of: casino royale

3/5 on 28 November 2006

Didn't go for the new tone. It wouldn't be fun to have this Bond's life, for one thing. There were frequently secrets from the audience (Bond knows M's name but won't tell us), and what few gadgets there were didn't even get used the way the movie tried to. ([More of casino royale... at Revyu.com](#))

Figure 24. Example output: detail pages such as this show items the information source has reviewed that relate to the topic of the query.

7.5.2. Ranking by Trust Factors

The first step in the process of assembling results to a query is to identify all individuals who have experience of the topic(s) of the query. The experience factor was not identified as exceptionally influential in the study of how people decide who to ask for recommendations, however it would appear to be more neutral than expertise in not being oriented towards tasks with particular characteristics (e.g. high criticality tasks).

For this reason, combined with the fact that experience metrics have greater coverage of topics than expertise, experience metrics are taken as the baseline in building sets of

search results. This is akin to a source identification process whereby those who may have relevant knowledge are selected from a larger pool which may then be further refined, and bears some likeness to McDonald and Ackerman's (1998) expertise identification stage.

By default search results are ranked by the experience of individuals with regard to the topic of the query, and presented to the user. However, as demonstrated by the results presented in Chapter 4, the role of trust in information-seeking is not constant but varied and situational, depending on characteristics of the task such as criticality and subjectivity. Therefore, mechanisms are required in Hoonoh to enable results to be ranked by different trust factors in a way that is sensitive to these task characteristics.

Two different approaches to this issue were considered. The first was to classify all topics in the system according to measures of criticality and subjectivity (the factors found to influence which trust factors were attended to). This would then allow algorithms running behind the scenes to select and employ appropriate combinations of trust factors before results were presented to the user. The second approach considered was to allow users to select the factors used to weight results according to their perception of task characteristics, and vary these in order to refine search results provided by the system.

The first of these options was deemed to be unnecessarily restrictive, potentially ignoring individual differences in perceptions of task criticality and subjective, and limiting the flexibility of the system to adapt to new topics. Consequently, users are able to vary the use of different trust factors in ranking of potential information sources by clicking on

links presented above search results, with the option to add expertise or affinity metrics to the existing experience based ranking (Figure 25).

Giving users the freedom to select the factors by which results are weighted is likely to increase the transparency of the system, which may in turn positively impact user acceptance (Herlocker, Konstan et al., 2000). However, one area in which transparency may need to be improved in the future is in communicating to users the means by which the trust metrics are generated, and therefore what action they may take to refine these.

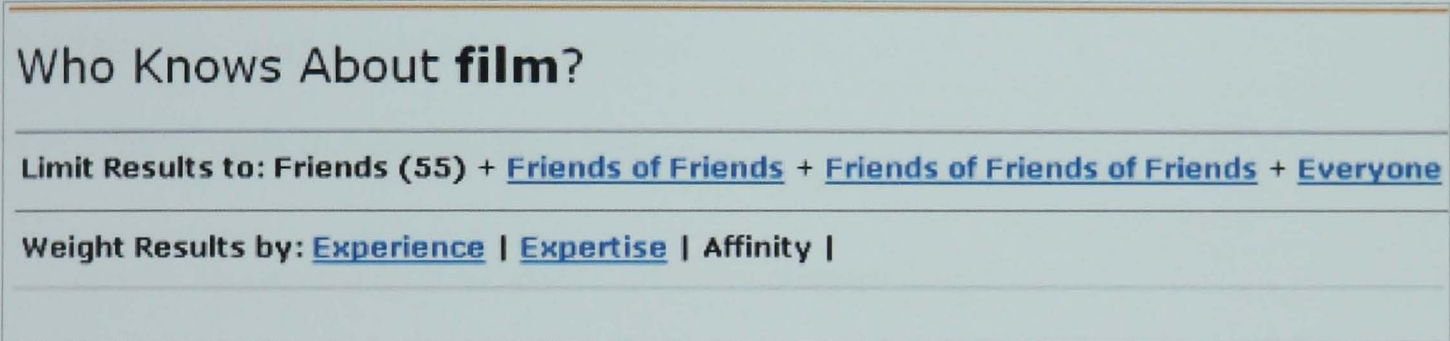


Figure 25. Weighting options in the Hoonoh results interface

Where additional ranking by expertise or affinity is requested, the expertise or affinity metric for each individual featured in the results is multiplied by a weighting factor and added to the experience score. Results are then ranked by this combined score. Each factor has a weight of 1 by default as a clear case for differential weighting of factors has yet to emerge; however, the system has this capability if it later proves necessary.

7.5.3. Filtering by Social Network

Users of Hoonoh are treated as anonymous by default, and can use the site and rank results by experience or expertise without logging in, as shown in Figure 26.

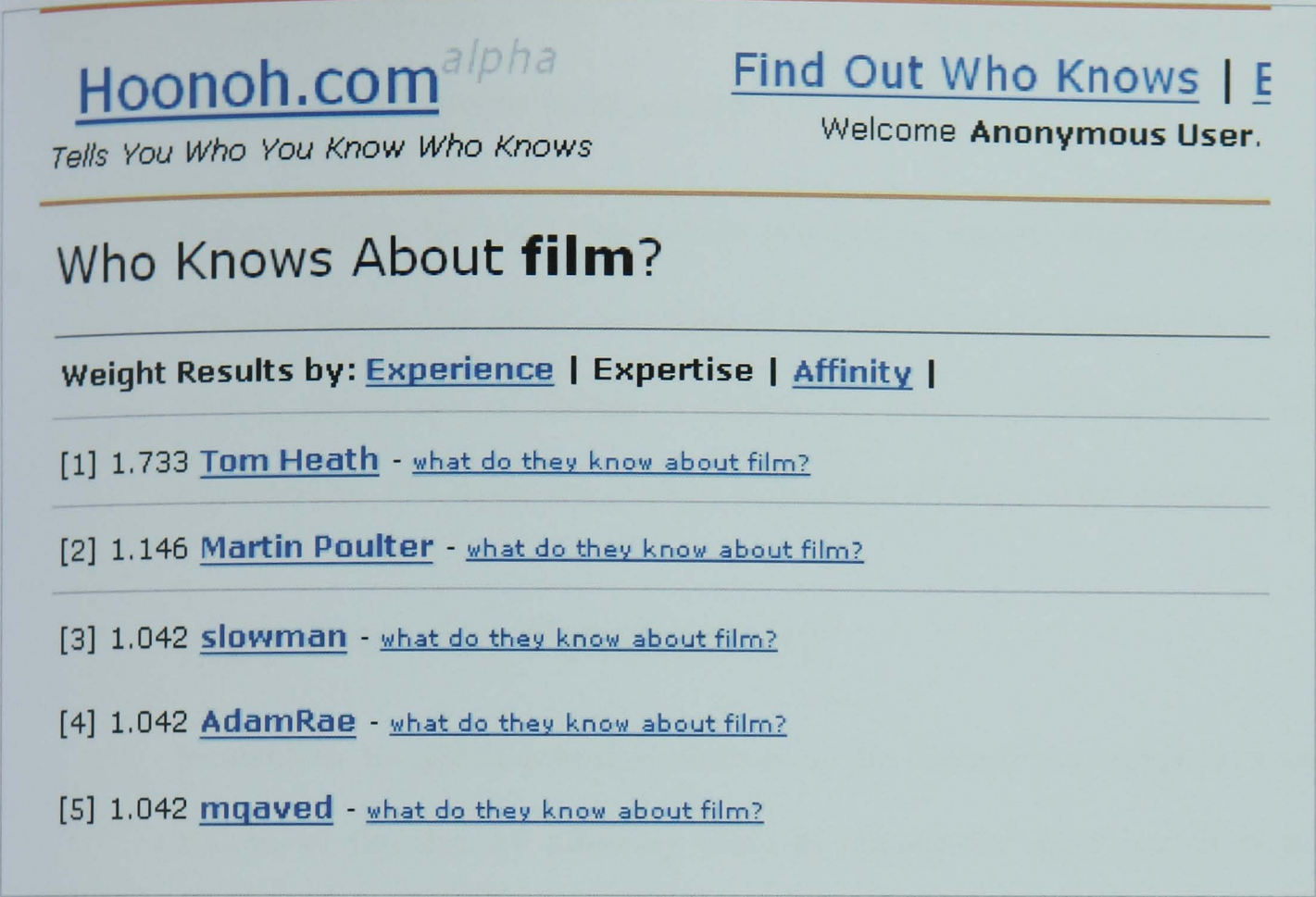


Figure 26. Example output: weighted results for anonymous users

Those who do choose to log in can also rank results by affinity (being a person → person relationship affinity-based ranking is dependent on knowing the user's email address).

Whilst ranking results according to the trust metrics generated by the Hoonoh algorithms may be useful in itself, the core functionality of Hoonoh is the ability to limit the individuals returned in search results to those within one's social network. This allows the exclusion of any items that may be relevant to the topic of a query but have not been recommended by known people. This provides a personalised view on the information-seeking process that is not available in existing Web search engines.

This functionality also has the potential to render spam reviews (i.e. those provided by individuals who have a vested or commercial interest in writing a good review) meaningless. As the user can determine who is listed in the social network that shapes the results they see, spurious entries from unknown individuals are automatically filtered out,

or known individuals who create biased or otherwise undesirable reviews can be excluded from contribution to search results

In cases where the user's immediate network of known individuals does not provide adequate results to a query, the scope of the search can be expanded to include friends of friends, and friends of friends of friends (i.e. those 2 or 3 hops away in the network, respectively). The scope can even be widened to all users of the system if desired.

7.5.4. Browsing Functionality

In addition to search-oriented interaction, the system also supports a general 'social awareness' function by allowing users to browse for other people in a topic-centric fashion (Figure 27), and for topics in a people-centric fashion (Figure 28).

Most Known-About Topics

[007](#) (8) [accommodation](#) (9) [aston-martin](#) (8) [bar](#) (11) [bb](#) (9) [beer](#) (14)
[bond](#) (8) [book](#) (12) [budva](#) (176) [cafe](#) (11) [camera](#) (8) [casino](#) (8)
[casino-royale](#) (8) [community-website](#) (8) [conference](#) (176)
[daniel-craig](#) (8) [dvd](#) (9) [eating-out](#) (15) [english](#) (8) [eswc2006](#) (176)
[film](#) (25) [food](#) (33) [friends](#) (9) [gambling](#) (8) [hollywood](#) (8) [james](#) (8)
[james-bond](#) (8) [london](#) (12) [milton-keynes](#) (21) [montenegro](#)
(176) [movie](#) (16) [movies](#) (15) [music](#) (17) [poker](#) (8) [pub](#) (15)
[restaurant](#) (38) [review](#) (13) [royal](#) (8) [semantic-web](#) (181) [tags](#)

Figure 27. The 'topics' area of the site allows users to browse for potential information sources by topic

Browse People

Your Network (55)

[Adrian Stevenson](#) [AdrianStevenson](#) [AlanJ](#) [Arthur Stutt](#) [Ben Lund](#) [Bertrand Sereno](#) [Bhagesh Sachania](#) [Brian Kelly](#) [Chris Mitchell](#) [Craig McKenzie](#) [Crash](#) [Damian Steer](#) [Damian Steer](#) [Dan Brickley](#) [Danni](#) [Dnyanesh](#) [Dnyanesh Rajpathak](#) [Enrico Motta](#) [Fin](#) [Geri](#) [Gregory Williams](#) [Ivan J Stevenson](#) [Jen Chambers](#) [Jianhan Zhu](#) [John Domingue](#) [Marc Eisenstadt](#) [Marian Petre](#) [Mark C Martin Dzbor](#) [Martin Poulter](#) [Michele Pasin](#) [Paul Hollands](#) [Sam Chapman](#) [Simon Buckingham-S](#) [Timo Hannay](#) [Timo Hannay](#) [cancer](#) [castagna](#) [cyqri](#) [danja](#) [domenico79](#) [drewp](#) [glittrgirl](#) [gromg](#) [hockeyshooter](#) [inez](#) [jccq](#) [kasei](#) [kidehen](#) [kjwa](#) [leobard](#) [magicrebirth](#) [xcv](#)

Top 30 People (By Topics Known About)

[Tom Heath](#) (YOU) (836 topics) [Martin Poulter](#) (151 topics) [hockeyshoo](#) (131 topics) [Paddy](#) (106 topics) [teddypolar](#) (93 topics) [drewp](#) (85 topics) [sofia](#) (84 topics) [mqaved](#) (78 topics) [Crash](#) (67 topics) [magicrebirth](#) (64 topics) [vladtn](#) (49 topics) [glittrgirl](#) (46 topics) [smonroe](#) (46 topics) [Mark](#) (42 topics) [Sanyukta](#) (39 topics) [redwards](#) (35 topics) [Fin](#) (34 topics) [Dnyanesh](#) (31 topics) [Aneta](#) (30 topics) [Kake](#) (29 topics) [AlexLittle](#) (29 topics) [LordByron](#) (28 topics) [jccq](#) (27 topics) [AdamRae](#) (26 topics) [bouquet](#) (25 topics) [xcv](#) (23 topics) [slowman](#) (22 topics) [Paul](#) (21 topics) [AdrianStevenson](#) (21 topics) [can](#)

Figure 28. The 'people' area of the site allows users to browse the topics in which specific individuals have experience or expertise. Logged in users will be shown people who the system knows to be in their network.

7.5.5. Queries for Multiple Topics

Users are able to enter search queries into Hoonoh that contain multiple topics, e.g. 'restaurant madrid'; up to four separate topics can be entered (any above this number are ignored for performance reasons). This capability enables users to enter queries for more complex topics that may not be captured by just one tag. In such cases separate result sets are computed for each term individually, and these are then summed, meaning that multiple topics are effectively treated as being part of a 'Boolean OR' query. One consequence of this is that someone with a high experience score for one topic but a low experience score for another may rank relatively highly in the combined results, and even equally to someone with moderate experience of both domains.

Whilst this could be seen as providing misleading results (the source may know nothing of one of the topics), a decision was made to adopt an open world view, on the basis that the person may have knowledge that is not captured in the system. Furthermore, it was decided that treating multiple topics as part of a 'Boolean AND' query may result in few or no results being returned for queries in which one topic is sparsely represented in the Hoonoh data set.

One additional factor that mitigates against any problems caused by this design decision is that someone with a high experience score for all topics in a query will be ranked higher than another person who has a high experience score in just one topic, maintaining the relevance of the results.

7.5.6. System Implementation

Figure 29 below reproduces the upper portion of Figure 5 from Chapter 5 and shows the architecture of the Hoonoh system. Trust metrics generated by the Hoonoh algorithms implemented in the Trust Computation Subsystem (see Figure 5) populate the Hoonoh triplestore. In common with Revyu, the Hoonoh triplestore is based on a denormalised MySQL database that follows the RAP database schema referred to in Section 6.8.

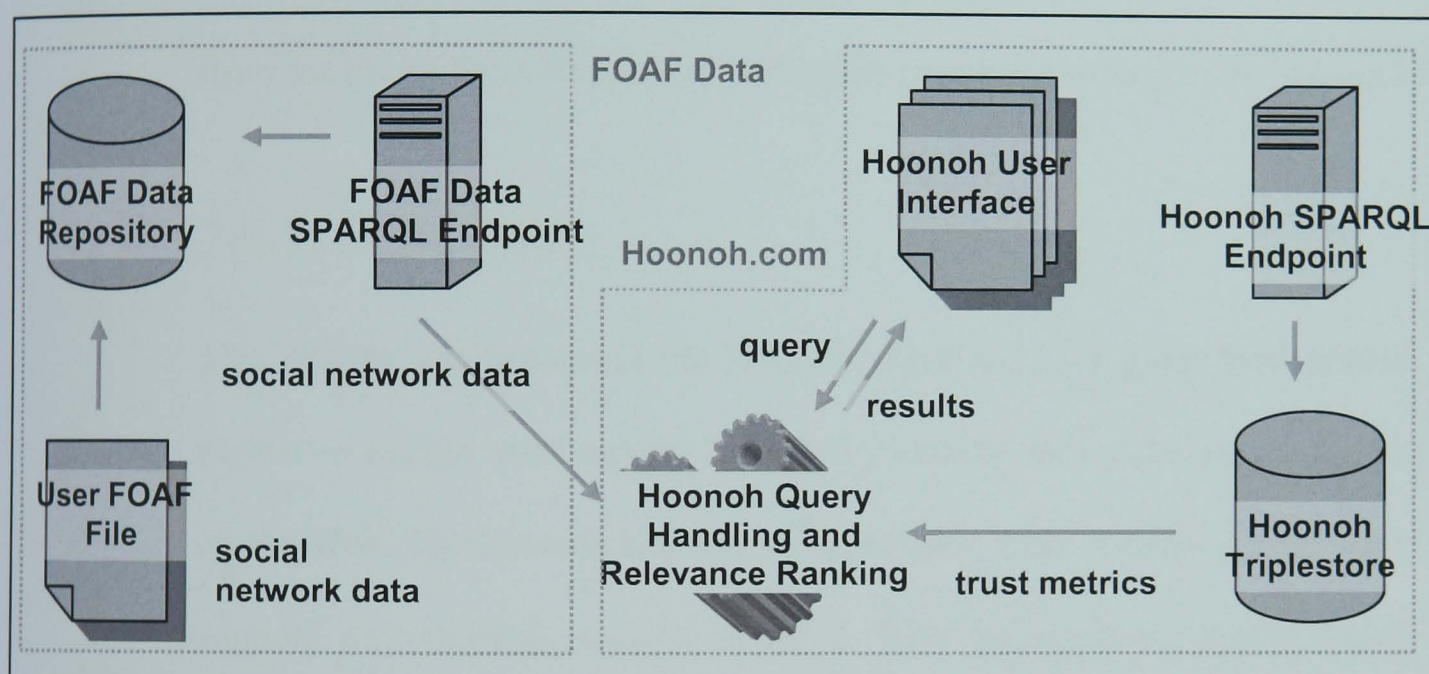


Figure 29. Hoonoh system architecture

For the reasons outlined in Section 5.4, the social network data used in this research comes in the form of FOAF files. In order to provide a data set with which to demonstrate Hoonoh, more than 70,000 RDF documents primarily containing FOAF data were crawled from the Web, producing a data set of over 6 million RDF triples. In addition to `foaf:knows` triples describing social network connections between individuals, this RDF includes additional generic data about individuals (such as names, addresses of home pages, etc.) that is not otherwise stored in Hoonoh. These 6 million triples were imported into a Talis Platform (Leavesley and Davis, 2007) store to form the

FOAF Data Repository shown in Figure 29 above and provide SPARQL query capabilities over this data set.

User queries received via the Hoonoh Web interface are passed to the 'Query Handling and Relevance Ranking' engine which is implemented in PHP and runs on a regular Apache Web server. This engine is responsible for identifying from the Hoonoh triplestore individuals who are potential sources of information on the topic(s) for which the user has searched. The engine combines this information with social network data from the FOAF Data Repository and ranks results according to the selected trust factor.

7.5.7. Summary

This chapter has presented the Hoonoh algorithms for generating person \rightarrow topic and person \rightarrow person trust metrics from Revyu rating data and a range of other data sources on the Web. The Hoonoh site demonstrates how these metrics can be used in a system to support personalised, socially-oriented Web information-seeking based on word-of-mouth recommendation principles. This demonstrates that implementation of such systems is feasible at a technical level. The following chapter presents an evaluation of the effectiveness of the Hoonoh algorithms, and by extension the Hoonoh system as a whole.

8. Evaluation

8.1. Introduction

Hoonoh, described in Chapter 7, was implemented as a test of the principles and approaches developed in this research. Having implemented the system an evaluation was carried out to achieve two things: 1) to assess the ability of the Hoonoh system as a whole (algorithms plus ranking engine) to generate data that predicts the members of their social network that individuals would choose as information and recommendation sources in different scenarios; and 2) to determine whether meaningful results could be produced with a minimal level of data input.

Herlocker, Konstan, Terveen et al. (2004) discuss methods for evaluating collaborative filtering recommender systems. One such technique that has been widely used involves withholding a portion of the rating data provided as input to the system, and using this to assess the ability of the system to accurately predict these ratings.

Evaluating Hoonoh requires a slightly different approach, primarily because the system is oriented towards producing output in the form of ranked lists of possible information sources, not towards predicting ratings on items. This subtle but crucial distinction means that an approach based on comparing ranked lists of results, and more akin to how one may evaluate a search engine, would be more appropriate.

Therefore, in this evaluation the efficacy of the system will be measured as the extent to which the sources it selects match the sources that human users would choose for the same task.

8.2. Design

In summary, this evaluation study used the Hoonoh algorithms to generate trust metrics for a group of participants, based on reviews they had created in Revyu. These trust metrics were used by the Hoonoh system to rank participants on particular topics and according to different trust factors. The rankings generated by the system were then compared to participants' own reports of how they would rank the other participants as information sources for the topics in question.

The evaluation was designed to address Research Question 6:

'If such systems can be implemented, how do they perform relative to human performance of equivalent tasks?'

This question can be rephrased in more operational terms as:

'How does output from the Hoonoh system, based on trust metrics generated by the Hoonoh algorithms, compare to participants' reports of the members of their social network they would use as information sources in recommendation-seeking scenarios?'

Based on this rephrased question, the primary hypotheses examined were:

- H_0 : there will be no correlation between participants ranking of potential information sources and the ranking of potential information sources produced by the Hoonoh algorithms.

- H_1 : there will be a significant positive correlation between participants ranking of potential information sources and the ranking of potential information sources produced by the Hoonoh algorithms.

8.3. *Method*

8.3.1. Participants

The sample used in the evaluation consisted of 17 participants opportunistically sampled from among staff and research students of the Knowledge Media Institute at The Open University. It was important that the sample had a number of characteristics:

1. A reasonable degree of 'knownness' between individuals was required in the sample, as awareness of another's knowledge is a prerequisite for assessing them as a potential source.
2. A reasonable (but not excessive) degree of overlap was required in participants' interests and areas of knowledge.

Therefore this department was chosen as it is small enough in numbers to ensure that participants are highly likely to know each other. It is also large enough to provide a population in which there are likely to be reasonably divergent patterns of interpersonal affiliations and social ties (i.e. social ties are unlikely to be uniformly distributed within the sample), which should allow for any effects of affinity to be identified. In contrast for example, a tightly-knit social group of close friends may be too strongly tied to allow for significant variation in source preference.

A degree of overlap in domain knowledge was necessary in order to identify a number of common topics to use as sample domains in the evaluation. It was important that all participants could feasibly be considered as information sources regarding the domain in question, to allow the analysis to be sensitive to variations in expertise and affinity as well as just experience.

For these reasons, participants in the evaluation were mostly new or existing but relatively light users of Revyu. No attempt was made to specifically recruit existing heavy users of the site to take part in the evaluation. Whilst this would have been beneficial in terms of greater number of reviews on which to base rankings, the degree of 'knownness' between these individuals was deemed too low to be useful in the evaluation. Furthermore, the overlap in domain interests and reviewed items was considered insufficient for these purposes.

Two domains were selected in which Hoonoh would be evaluated: restaurants in Milton Keynes, and professional-quality cameras. Restaurants were chosen as a relatively subjective domain in which affinity may play a role in source selection, whilst professional camera equipment was chosen as a relatively critical domain in which expertise may be beneficial.

8.3.2. Procedure

All 17 participants were asked to use Revyu to create six reviews; two each from the following three categories. Explanatory notes given to participants are shown in brackets:

- **Things in Milton Keynes** (This might be a pub, restaurant, shop, service or leisure activity. Have a look at <http://revyu.com/tags/milton-keynes> for some ideas of things to review.)
- **Consumer Electronics Items** (This might be a new camera, mobile phone, iPod or piece of computer hardware. If you are unsure what to review, think of the last electrical items you bought and consider reviewing those. If you just can not think of what to review in this category, review extra items in the next category (Items of Your Choice).)
- **Other Items of Your Choice** (This can be anything you like (although preferably not people!). A book, film, or music album can all be good choices, as can other things in Milton Keynes that you might not have reviewed under the first category, or things located in another town/city/country.)

The precise type of items to be reviewed was deliberately left loosely specified in order to allow personal interests to influence the results. For example, someone who enjoys eating out and therefore has extensive experience of restaurants may choose to review items of this type. In contrast, requiring someone with little interest in this domain to review such items may create somewhat artificial experience scores for a domain in which they have relatively little knowledge and would not otherwise choose to rate items.

The frequency of reviews created by each participant is shown in Table 6 below.

Participant	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
Reviews Created	6	6	6	5	18	7	6	10	6	10	11	6	4	6	12	6	8

Table 6. Frequency of reviews created by each participant in the evaluation

The higher frequencies shown in the table represent participants who were existing users of the site and had already contributed reviews prior to taking part in this evaluation. These individuals were invited to contribute additional reviews specifically for the evaluation, but not required to do so if their reviews already met the requirements of the analysis.

Not all participants provided a sufficient number of the required type of reviews to fully meet the criteria. For example, two participants ('C' and 'G') were only able to review one thing within Milton Keynes due to limited familiarity with the town. Participants F and P only reviewed one consumer electronics item, whilst D and M reviewed none at all. Participant M only reviewed things in Milton Keynes, omitting to review consumer electronics items. Despite these discrepancies, all participants were retained at this stage of the study in case further insights could be gained from participants with low numbers of reviews.

Only rating data from Revyu was used in the evaluation, in order to assess the quality of results derived from the core Hoonoh algorithms.

In Phase 2 of the study each participant was given a questionnaire that outlined two scenarios, reproduced below:

Scenario 1 - Restaurant Recommendations in Milton Keynes

'Imagine you are planning to take a friend out for a meal in Milton Keynes, and need a restaurant recommendation.'

Scenario 2 - Camera Recommendations

'Imagine you are looking to buy a new camera of professional quality, and need a recommendation about which model to buy.'

Below each scenario the questionnaire listed all other individuals taking part in the study in a random order, and the participant was asked to rank these individuals according to the order in which he or she would choose to ask them for recommendations on the topic of the scenario. For example, Participant A would be asked to list participants B to Q in the order in which each would be chosen as an information source in both the restaurants and cameras scenarios.

This approach was preferred over alternatives such as presenting participants with a ranked list of information sources for a topic and asking them to rate the accuracy of the list; such an approach was considered too vulnerable to confirmation biases (where participants selectively attend to information that fits their existing schemas) and response biases (where participants give responses they believe the experimenter is seeking).

To accommodate cases where an individual was unable to rank others due to insufficient knowledge of other participants, the following instructions were provided alongside those explaining the ranking procedure:

'If you genuinely don't know some of the people on the list then leave them unranked, however do please try and rank as many of the names as possible.'

Finally, as described in Section 7.2.4, a snapshot of the entire Revyu data set was taken that included all reviews created by participants in the study. The Hoonoh algorithms were executed using this snapshot, to generate trust metrics for all Revyu reviewers.

A social network definition was created for each participant using the FOAF vocabulary. This simply stated that each participant `foaf:knows` all other individuals taking part in the study. Other reviewers who a participant may know, and friends of friends, were excluded from this FOAF file, as evaluation data was not being collected from these individuals. In summary, the network was limited to a radius of one hop and only those individuals who were participating in the study.

8.4. Analysis

Trust metrics that related to the evaluation participants were used in combination with the FOAF files generated for the study as input to the Hoonoh Relevance Ranking engine, to produce ranked lists of potential information sources from among the other participants in the study. Several such lists were produced for each person and for the topic of each scenario used in the evaluation, based on ranking of results by different factors and the application of different weights to these factors.

Ranked lists of results were generated based on the following combinations of factors and weights: experience alone; experience and expertise weighted equally; experience and affinity weighted equally; experience and expertise with expertise given twice the weight of experience; experience and affinity with affinity given twice the weight of experience. These combinations are summarised in Table 7:

	Ranking 1	Ranking 2	Ranking 3	Ranking 4	Ranking 5
Experience Weight	1	1	1	1	1
Expertise Weight	0	1	2	0	0
Affinity Weight	0	0	0	1	2

Table 7. Combinations of trust factors and weightings used to generate ranked results

Participants had been asked to review two items of their choice from the domain of consumer electronics items. However, the specific topic of Scenario 2 in the evaluation was camera equipment. As discussed previously, this loose specification of items to be reviewed was deliberate in order to allow areas of personal interests and knowledge to be reflected in the results.

Eight participants were identified as having experience of the topic 'camera' based on the trust metrics derived from Revyu reviews. Whilst small, this number can still be subjected to statistical analysis.

However, further examination of the output from the ranking engine revealed a very high number of tied ranks in the results for the camera scenario; i.e. several people received the same score and therefore share a rank in the results. Seven of the eight individuals listed in results were tied when only experience and/or expertise were used for ranking. In these cases results were completely determined by experience scores; a lack of convergence among participants in what was reviewed in the 'consumer electronics' domain appears to have prevented the generation of expertise metrics related to the 'camera' topic. As a consequence expertise scores did not add any variation to overall scores for Rankings 2 and 3, which could have reduced the number of tied ranks in the results for this scenario.

As a result of this high number of tied ranks, this data set could not be subjected to the correlational analysis carried out on the restaurant scenario data (described below). However, some interesting trends are apparent: 13 of a possible 16 participants ranked the same individual in first place as a source of information and recommendations related to cameras. Three other individuals also consistently appear in the top three positions in participants' ranking of others in the camera domain. Possible interpretations of this outcome are discussed below.

Output from the Hoonoh ranking engine for the restaurant domain contained few tied ranks. This enabled Spearman rank correlation coefficients to be computed for each participant in the study for the restaurant scenario, with the exception of participant C who was not able to rank a sufficiently high number of people to warrant statistical analysis (due to not knowing enough of the other participants well enough to make a judgement).

All 17 participants in the evaluation had experience scores in the 'restaurant' or 'milton keynes' domains, or both; therefore all participants appeared in the system output for the restaurant scenario. The rankings from the Hoonoh ranking engine were compared to the responses provided by participants when ranking their preferences for other participants as recommendation sources in the restaurant scenario. Where a participant had ranked fewer than 16 of the participants, those individuals who had not been ranked by the participant were removed from the results used in the analysis. Any remaining tied ranks were then resolved by averaging, and a Spearman rank correlational analysis was carried out on the data. Following the guidance of Howell (2002) the formula for calculating the Pearson correlation coefficient was applied to the ranked data used in this evaluation, in

order to produce Spearman coefficients resilient to the small number of tied ranks in the restaurant data.

8.5. Results

Table 8 shows the coefficients (r) demonstrating the Spearman rank correlation between the responses of each participant and the Hoonoh system output (using the various ranking combinations defined above). The column titled ' n ' shows the number of participants who were ranked by each participant. Values of ' r ' marked in bold are significant at the 0.05 alpha level for one-tailed tests. Critical values are taken from Ramsey (1989).

Participant	<i>n</i>	Critical Value (<i>p</i> = 0.05)	Ranking 1 <i>r</i>	Ranking 2 <i>r</i>	Ranking 3 <i>r</i>	Ranking 4 <i>r</i>	Ranking 5 <i>r</i>
A	8	0.643	0.443	-0.024	0.190	0.310	0.310
B	7	0.714	0.364	0.432	0.396	0.364	0.364
D	14	0.464	0.288	0.101	0.119	0.232	0.232
E	9	0.600	0.661	0.717*	0.65	0.661	0.661
F	16	0.429	0.586	0.015	0.012	0.586	0.586
G	10	0.564	0.200	-0.188	-0.164	0.200	0.200
H	16	0.429	0.049	0.320	0.346	-0.021	-0.021
I	7	0.714	-0.036	0.071	0.071	-0.036	-0.036
J	12	0.503	0.223	-0.301	-0.252	0.223	0.223
K	16	0.429	0.211	-0.015	-0.018	0.277	0.277
L	16	0.429	0.296	0.075	0.060	0.296	0.296
M	16	0.429	0.467	0.180	0.159	0.357	0.348
N	13	0.484	0.590	0.209	0.132	0.595	0.595
O	11	0.536	-0.083	0.351	0.405	-0.032	-0.032
P	9	0.600	0.153	0.067	0.067	0.209	0.209
Q	16	0.429	0.361	0.243	0.231	0.361	0.361

Table 8. Spearman rank coefficients of correlation between participant responses and Hoonoh output with each ranking combination. Entries in bold are significant at the 0.05 alpha level.⁴⁰

Overall, ranking based on experience alone (Ranking 1) produced the greatest number of statistically significant results. Therefore in these cases (and others showing significant relationships under different rankings) the null hypothesis that there would be no correlation between participants' responses and system output should be rejected. In the remaining cases the null hypothesis of no correlation should be accepted.

The lowest number of statistically significant results was achieved when ranking was based on a combination of expertise and experience, at either level of weighting (Rankings 2 and 3). In one case (Participant F) ranking based on experience and expertise

⁴⁰ The entry marked * is significant at the 0.025 alpha level. Participant C is excluded due to a low value of *n* preventing statistical analysis.

reduced an existing correlation based on experience scores from a statistically significant 0.586 to almost zero (0.015), and in another two (Participants M and N) a significant correlation was reduced to below the level of significance.

In another case (Participant E) inclusion of expertise at a weighting equal to that of experience (Ranking 2) increased an already significant correlation slightly (by 0.056), producing the single highest correlation found in the study ($r = 0.717$, significant at the 0.025 level). However, when expertise was given twice the weighting of experience (Ranking 3) the correlation was reduced very slightly below the original level in Ranking 1.

It appears that the relatively low number of reviews on which expertise scores for the 'restaurant' and 'milton keynes' domains are based may have skewed these scores, leading to the reduction in otherwise significant correlations seen with Rankings 2 and 3. A larger and more convergent data set on which to base expertise scores is likely to resolve this issue.

Affinity had little effect in most cases, irrespective of the weight it was given. In one case (Participant N) it produced a very slight positive increase in an already significant correlation; in another case (Participant M) lowered an existing correlation based on experience scores alone to below the level of statistical significance.

The very limited impact of affinity on the correlations found between participants' responses and the rankings produced by Hoonoh can be attributed to the low numbers of affinity ratings between participants in the study. Only nine of the 17 participants had any affinity relationships to others in the study, of which there were only 16 in total (a mean of less than one affinity relationship per participant). This does not mean that affinities do

not exist between those who took part, but instead reflects the low overlap in items reviewed by participants in the study, which in turn limited the computation of affinity relationships.

8.6. *Discussion*

Drawing firm conclusions from the results of the evaluation is not straightforward. The existence of a number of significant correlations demonstrates that in some cases Hoonoh succeeds in replicating the source selection decisions of participants. However in other cases no correlation is found and reasons for this discrepancy are not immediately apparent.

One possible explanation is that the Hoonoh algorithms have some innate characteristics that lead to generation of metrics that are more attuned to certain peoples' source selection processes, or more aligned to certain peoples' knowledge of what other people know; i.e. is there some bias inherent in the algorithms that happens to fit with these people?

Alternatively the non-significant correlations may reflect inaccurate judgements by participants about the domain knowledge of other people. This may be based on a lack of exposure to what other people know and thus poor ranking of potential sources, whilst the system-generated metrics may in fact be a more accurate reflection of individuals' domain experience. One factor arguing against this conclusion is the observation that the significant correlations found in response to Ranking 1 are evenly distributed between participants who are long-term members of the department and others who were relatively new at the time the evaluation was carried out. One would expect long-term

members to have a better awareness of the knowledge of other participants in the sample, and thus more accurate rankings if assessment was based purely on experience.

One observation worth noting is that the significant correlations can be explained primarily as a function of experience scores; in all cases the role of expertise or affinity is minor or in some cases even negative. The results for Participant E do show a slight increase in correlation when expertise is introduced and weighted equally with experience (Ranking 2), however this increase is small and unlikely to be statistically significant. It is the only case in which the addition of expertise increases an existing correlation based on experience metrics.

Consequently, perhaps the most viable explanation for the variation in results across the participants is that some people more than others may naturally place greater emphasis on factors such as expertise or affinity when seeking information, or were more influenced by these factors during this evaluation.

If this is the case, then the low numbers of expertise and affinity scores that could be derived from the data set used in this study would have limited the ability of the system to match participants' rankings, which would in turn explain the number of significant correlations obtained.

Therefore at this stage the precise effects that expertise and affinity metrics may have on improving results from the Hoonoh system is not apparent. Future work may provide insights into this issue, however considerably larger amounts of data, exhibiting a higher degree of item overlap, will be required in order to make these assessments. Section 9.3.6 briefly discusses some of the challenges associated with acquiring benchmark data sets for evaluation of systems of this type.

Whilst the results for the camera scenario were not amenable to statistical analysis, they do provide some support for the importance of expertise in ranking Hoonoh results in some scenarios. Eleven out of 16 participants chose the same first source for information and recommendations about cameras; three other individuals were also heavily represented in the top three positions. The strongest explanation for this consistency is that the participant frequently listed as first choice occupies a role of an expert within the group with regards to this topic. An alternative explanation that all participants have very strong affinity to this individual is not supported by the data from the restaurant scenario.

Cross and Borgatti (2004) found evidence for the existence of major nodes or hubs in information-seeking networks, individuals who were frequently sought out as sources of information. The data from the camera scenario may indicate a similar structural feature in the group of participants with regards to information about cameras.

The evaluation results provide some tentative evidence that Hoonoh may be able to produce results of some value even with minimal bootstrapping by the user (in the form of writing reviews, for example). This goes some way to addressing the second aim of the evaluation, that of assessing the minimum level of data input required by users in order to achieve meaningful results.

In some cases the simple answer to the question '*how much data must users provide to benefit from the system?*' would appear to be 'zero'. All significant correlations found in the study would have been significant through experience metrics alone (and some were significant only when experience metrics alone were used). Therefore any one participant could have achieved the same results for herself as long as all others supplied sufficient

data that allowed derivation of metrics about their own experience and expertise (both parties would need to supply data to generate affinity scores).

On this basis there would appear to be a strong case for integrating considerably larger amounts of background data into the system to boost coverage of experience metrics and therefore increase the utility of the system for a greater number of users. Chapter 9 discusses ways in which this may be achieved.

The '*no bootstrapping required*' conclusion raises the issue of whether users of a system like Hoonoh would attempt to 'freeload', relying on others to invest the effort in providing ratings and other data but benefiting from this through better search results. Perhaps the best safeguard against such a scenario is the potential importance of affinity metrics as the volume of information in the system increases. In this case, and particularly in subjective domains, experience and expertise metrics may prove insufficient without affinity, requiring users to engage in providing data in order to enable affinity metrics to be generated.

An additional interesting feature of the data is the variation in numbers of people ranked by participants in their responses, ranging from 4 to 16 in the restaurant scenario and from 2 to 16 in the camera scenario. This may indicate a number of things: individuals may differ in their willingness to ask other people for advice, and therefore more cautious individuals may list fewer potential sources; it may indicate other factors such as variation in 'knowingness' within the sample, i.e. long established members of the department may know (and therefore be willing to ask advice of) a greater number of other people; or it may reflect issues such as variations in the perceived availability of others, a factor identified by Cross and Borgatti (2004).

Alternatively, another explanation may exist. In five cases participants varied individually across the two scenarios in how many people they ranked as information sources. This may indicate that people's awareness of who knows what may vary across domains, and highlights another key feature of Hoonoh: informing users about members of their networks who have relevant knowledge, of which they may not be aware. Cross and Borgatti (2004) conclude that on balance previous literature suggests that 'knowing who knows' is "probably the single most important variable in knowledge seeking", supporting the need for this kind of functionality. Therefore, irrespective of the ranking of results Hoonoh can provide a social awareness function that might not otherwise be available.

9. Conclusions and Outlook

9.1. Summary of the Research

The research presented in this dissertation has examined how information-seeking on the Web can be better supported by harnessing the knowledge of trusted members of our social networks. In doing so I have sought to answer the following general question:

'To what extent can information- and recommendation-seeking within social networks be supported on the Web?'

In order to answer this main research question, six specific Research Questions were formulated:

1. How do people choose sources for information and recommendations from among members of their social network?
2. Which factors influence judgements about the appropriateness and trustworthiness of these information and recommendation sources?
3. How do the characteristics of the task being performed affect these judgements?
4. To what extent can general principles derived from answers to the previous questions be operationalised as computational algorithms that replicate the process of seeking information and recommendations from social networks?
5. How feasible is the implementation of user-oriented systems that exploit such algorithms?

6. If such systems can be implemented, how do they perform relative to human performance of equivalent tasks?

Literature reviewed in Chapter 2 demonstrates the importance of social networks in information- and recommendation-seeking; both as information sources, and as quality filters that reduce information overload and increase the relevance of information to our needs. Previous researchers (e.g. Kautz, Selman et al., 1997b, 1997a) have investigated related issues and attempted to develop systems that exploit these processes.

My approach to supporting these processes on the Web is described at a conceptual level in Chapter 3. This approach is significant in that it examines the source selection process from first principles and pursues this through algorithms and implemented systems which are subsequently evaluated. In doing so the research makes four major contributions.

9.2. Contributions of the Research

9.2.1. Contribution 1

In order to maximise the value and effectiveness of word-of-mouth recommendation it is important to select the most appropriate information sources. Existing literature has much to say on the matter, however this is mostly confined to either workplace settings or taste domains, as discussed in Chapter 4. The first three research questions address issues raised by the shortcomings of previous work on source selection in word-of-mouth information-seeking. It is in addressing these questions, and providing a richer understanding of the source selection process and at a more general level, that this research makes its first major contribution, presented in Chapter 4:

- An empirical study of decision making in recommendation-seeking identified five trust factors that influence the choice of information sources and their perceived trustworthiness. Variations were identified in how these factors are applied across situations with varying levels of criticality and subjectivity.

These findings provide a basis for systems that may support the source selection process across a range of different tasks. Those that are more critical in nature, and poorly served by current recommender approaches, may benefit greatly from the support of trusted social networks, especially where trust is defined in a task-appropriate fashion.

9.2.2. Contribution 2

Chapter 5 outlines the technical approach I adopted in order to address research questions 4-6.

Shortcomings were identified in the data available on the Web with which to investigate these questions. These shortcomings are outlined in Chapter 5, and resulted in the second major contribution of this research:

- Revyu, a live, public reviewing and rating Web site. The site is built on Semantic Web technologies to enable integration of review data with social networks, and easy reuse of the data in deriving word-of-mouth related trust metrics.

Providing review data that is more easily reusable has tangible technical benefits. It also opens review data up to a wider range of systems and service providers who may not otherwise have had access to such information. This may in turn lead to a greater number of systems that develop functionality based on reviews in order to better serve their users.

9.2.3. Contribution 3

In Chapter 7 I present the third major contribution of this work, which directly addressed the fourth research question:

- The Hoonoh algorithms are a set of three algorithms based on the empirical findings presented in Chapter 4. Using review data and additional background data sources on the Web the algorithms generate person \rightarrow topic and person \rightarrow person trust metrics that can be used in systems supporting word-of-mouth recommendation.

These algorithms represent an attempt to operationalise complex and subtle human decision-making processes. Section 9.3.3 considers the appropriateness of the credibility algorithm in more detail, whilst Section 9.3.6 discusses issues that may be raised by attempting to evaluate systems such as Hoonoh.

9.2.4. Contribution 4

Research Question 5 relates to the feasibility of implementing systems that exploit these algorithms and the metrics they generate. This question is answered to a significant extent by the fourth major contribution of my work, also presented in Chapter 7:

- Hoonoh, a live, public Web-based system supporting information-seeking by highlighting the knowledge held by members of peoples' social networks. The system takes a source-centric perspective on information-seeking, allowing users to search or browse for those who have knowledge of a particular topic and then rank them according to the trust metrics generated by the Hoonoh algorithms.

Whilst Hoonoh goes a long way towards answering Research Question 5, some feasibility issues do remain with implementing systems of this nature. These are discussed below.

Research Question 6 asks how such systems might perform as source selection applications, relative to human performance of equivalent tasks. This question is addressed by the evaluation presented in Chapter 8, which did demonstrate some statistically significant results. However, these were not consistent across the data set, suggesting there is ample room for improvement.

9.2.5. Summary

In relating the findings back to the original over-arching research question, it would appear that information- and recommendation-seeking processes are open to empirical examination, yielding results that can be formalised as algorithms in technical systems. The most pressing and immediate challenges would appear to relate to availability of data upon which these algorithms can operate, and enhancing the evaluation strategies with which the effectiveness of these algorithms can be assessed. Addressing these issues will help to ensure that the theoretical and technical basis that has been established for supporting Web-based information seeking in social networks is complemented by realisation of tangible user benefits.

9.3. *Limitations and Future Work*

This section will use a number of themes to shape discussion of some limitations in the existing research, alongside discussions of how these may be addressed by future work.

9.3.1. Availability of Data

The issue of available data has shaped many aspects of this research, and continues to be a limiting factor. A lack of review data that was readily available and in a form that enabled integration with social network data led to the creation of Revyu. Whilst Revyu has provided a substantial amount of data with which to test the ideas in this research, a significant increase in available data is required if the benefits of my approach are to be fully investigated. A number of approaches are being considered in order to address this.

Further Pre-population of Revyu

One approach to increasing the amount of available data in the system is the pre-population of Revyu with skeleton records describing things that people may wish to review, in order to attract potential reviewers to the site. Use of this technique with data from the Open Guide to Milton Keynes was described in Section 6.11.3.

This approach has been considered with a number of significant data sets, such as descriptions of roughly 12,000 films from *DBpedia* (Auer, Bizer et al., 2007) and 70,000 hotels from *Geonames*⁴¹. Being Semantic Web data sets, integration of Revyu with data from these sources would enable a number of linking opportunities that could greatly enhance the site at a user and data level.

However, initial investigations have identified a number of issues with this approach. The amount of data cleansing required with external data sources can be substantial, in order to address issues such as encoding of foreign characters and removal of bogus data

⁴¹ <http://www.geonames.org/>

generated by automated methods. Translation of the cleaned data into a format suitable for consumption can also be very resource intensive. This process involves taking the source data as input and generating new RDF graphs that are suitably structured for import into Revyu. Much of this can be achieved using SPARQL CONSTRUCT queries (Prud'hommeaux and Seaborne, 2007) for graph transformations; however except through use of property functions SPARQL does not provide string manipulation functions essential for this kind of data processing, such as when minting URIs. As a result, much of the processing must be carried out programmatically, which in turn increases the resource requirements.

Whilst this degree of data cleansing can be viewed as an acceptable one-off cost for static data sets (such as the ISWC+ASWC 2007 papers, as described in Section 6.11.3), the costs in terms of manual intervention and the robustness of such process make them questionable for data sets that are dynamic. For example, reprocessing *DBpedia* data following each new release as a means to update Revyu with data about new films is resource inefficient and likely to be unreliable. Applying the same principles to hotel or restaurant data introduces further data management issues, as workflows need to be established for removing records of establishments that cease to exist.

Pre-population techniques will continue to be investigated in ongoing development, however before significant additional resources are invested in this approach a detailed analysis is required into whether the number of reviews that pre-population yields above what is possible with organic growth justifies the cost.

Review Aggregation

One alternative approach to increasing the volume of data in Revyu is to perform aggregation of reviews from across multiple sources on the Web. This has been considered many times and always rejected for the reasons outlined in Chapter 5 related to reviewer identity and integration with social networks: populating Hoonoh based on anonymous review data would lead to a very poorly interconnected data set from a social point of view, as trust metrics would be generated for many people who could not be connected into a social network due to their anonymity.

Whilst the rationale for rejecting the aggregation approach remains valid, it may provide sufficient data (of sufficient quality) about reviewed items with which to populate Revyu fully automatically. This data may then attract further reviews through Revyu itself that would be suitable for populating Hoonoh. The major drawback to this approach would be the costs associated with matching records that originate from different data sources but refer to the same item.

Additional Data Sources

In addition to enhancements to Revyu, it is essential that future work investigates greater use of external data sets.

For example, one possible approach would be to extract significant keywords from the Web pages people provide when reviewing an item. This could give a broader indication of the item being reviewed and supply additional data for generating experience metrics.

The techniques used to integrate *Del.icio.us* bookmarking data with the Hoonoh algorithms could also be extended to other Web2.0 data sources. For example, *Flickr* may provide a particularly good basis for assessing people's experience of particular

locations or activities, as photos are likely to be tagged with a location name or words describing popular activities. Unfortunately issues of synonymy and particularly polysemy remain with this data source and may prove even more problematic: an individual may upload many photos of his favourite restaurant, the (fictional) 'Hawaii Bar and Grill' in London, tag the photos 'hawaii' and be assigned a high expertise score for this topic, without having visited the state of Hawaii.

Despite the wealth of data available from Web2.0 sources, further exploitation of arbitrary Semantic Web data is preferable to greater use of tagging or keyword data. As will be discussed below, Semantic Web data can provide richer information from which to determine people's domains of experience and expertise, and through the use of URIs is not subject to the limitations that stem from the syntactic nature of tagging.

The contents of peoples' FOAF files provide a potentially rich source of information about users' experience of particular topics. For example, where a user states in her FOAF file that she is `foaf:based_near` a particular location, it may be reasonable to conclude that she has some experience of that location, and consequently increase her experience score for this topic. Similarly, if an individual works for a particular organisation (expressed using the `foaf:workplaceHomepage` property) and the location of that organisation can also be determined from data on the Semantic Web, there may be a case for increasing the individual's experience score for that location.

Furthermore, the interests which people specify in their FOAF files may also provide a basis for cautious assertions about an individual's experience of particular domains. In addition a large overlap in stated interests between two individuals may provide weak evidence of an affinity relationship.

The major aspect of this approach that requires future research is the question of how to scale such techniques to the Semantic Web at large. Mining experience metrics based on analysis of a small number of hand-picked relationships described in RDF (such as `foaf:based_near`) is a feasible way to generate large amounts of data, but would in the process overlook many rich sources of data that happen to be described using different ontologies.

A reasoning infrastructure is required that allows operations of this sort to be carried out over truly arbitrary Semantic Web data, whilst also providing mechanisms for weighting the confidence in any derived trust metrics. Such an infrastructure must retain provenance information in order to demonstrate to users the sources from which these metrics were derived. Explaining to users exactly how the algorithms performed the derivation will become increasingly complex as the number of source data sets becomes vast, but remains an important feature in ensuring user acceptance (Sinha and Swearingen, 2002).

Mechanisms that may be used to generate or supplement trust metrics from arbitrary Semantic Web data are discussed in more detail in Section 9.3.4 below.

FOAF Data

The availability of FOAF data describing social networks has not been a limiting factor in this research *per se*, as sufficient data has been gathered to demonstrate how the system operates. However, a number of issues related to FOAF data must be addressed if the principles presented in this dissertation are to reach a mass audience, namely coverage, density and technical quality.

Despite a number of major social networking sites such as *LiveJournal* and *MyOpera* exposing FOAF descriptions of site users, the coverage provided by the resulting data set is still very small relative to the number of Web users in total.

Hoonoh does have a number of features that enable the system to function and provide user benefits without any knowledge of the user's social network; for example, ranking of results by any weighting factor is possible as long as the user has identified himself with an email address and, in the case of affinity, provided some data from which trust relationships to other users can be computed. However, this does not convey the benefits of using known individuals as recommendation sources; furthermore, as the number of people represented in the system inevitably grows, the role of filtering by social network will become increasingly important.

For Hoonoh to reach widespread adoption there must be some means to increase the coverage provided by FOAF descriptions of social network data. Recent social network interoperability initiatives such as *OpenSocial*⁴² show promise in this direction. However, despite the popularity of sites such as *Facebook*, the users of such systems only represent some fraction of the population. It remains to be seen whether the majority of people will be prepared to make social network information available on the Web in a form that is usable by services such as Hoonoh, and (potential) users' views on this issue will need to be taken into account. Encryption or partial encryption of RDF graphs (such as in the work by Giereth, 2005) may be one solution to this issue.

⁴² <http://code.google.com/apis/opensocial/>

The FOAF data crawled from the Web to support this research also revealed a surprisingly low number of `foaf:knows` relationships in most FOAF files. These have not been formally analysed, however the subjective impression was that the resulting social graph as represented by the FOAF data was of relatively low density (i.e. low numbers of interconnections between individuals). Consequently a concerted effort is required at the community level to address this issue and increase the overall density of the FOAF-based social graph, as this could limit the utility of a system like Hoonoh by restricting the scope of results returned to those from an artificially small social network. One significant requirement of this process is the replacement of Blank Nodes (Klyne and Carroll, 2004) in FOAF files with URIs, enabling identities to be reconciled across many sources without having to resort to 'smushing' of data based on *Inverse Functional Properties* (Dean and Schreiber, 2004) such as `foaf:mbox_sha1sum`.

However, even where individuals are identified by URIs, there will be an increasing need for applications to perform identity reconciliation when handling data aggregated from across the Web. For example, each service that exposes FOAF data will likely mint a service-specific URI for each user. This is to ensure that it is possible to look up a user and retrieve information about them held by that service, subject to being authorised to do so. This would not be possible if services all used URIs in third party namespaces.

The result of this process is that users may already have multiple URIs on the Web, and will likely acquire more over time. Even if the user chooses to connect all these identifiers using `owl:sameAs` statements, applications must be able to smush the data from distributed sources to construct an integrated profile of the user. At present the tools and technical infrastructure to perform such operations is are not widely available, but these will become increasingly important for applications such as Hoonoh.

The effort involved in maintaining one's personal FOAF file may be one major reason for the low density perceived in the data used in this research. Compared to an application such as *Facebook*, adding friends to one's FOAF file is a laborious process, and may also account for the poor data quality observed in some of the data set, both at a semantic and syntactic level. Significant portions of the data were rejected from the live FOAF data store due to errors in RDF/XML syntax, whilst experience with the use of public FOAF data in Revyu suggests that many FOAF files do not use classes and properties such as `foaf:PersonalProfileDocument` and `foaf:primaryTopic` which indicate who is the subject of a particular file and which greatly simplify the consumption of this data.

9.3.2. Scale

In addition to questions of what is needed to ensure Hoonoh is useful to the wider population, the issue of scale impacts at algorithmic and technical levels.

Whilst the foundation of Hoonoh in fundamental research on trust in human recommendation-seeking helps ensure it is based on sound principles, it is not apparent at this stage how well the system will scale from a user perspective. A point may be reached where the system contains a volume of knowledge that renders even the relevance mechanisms developed here unworkable. Such a scenario is analogous to how search engine algorithms have had to be modified as the Web has grown in size and complexity.

The current technical implementation of the algorithms performs very well on moderately sized data sets from Revyu; as described in Chapter 7 computation of several thousand usage, credibility and affinity relationships by the Trust Computation Subsystem is possible in less than 40 seconds. Supplementing these usage scores with data from *Del.icio.us* and arbitrary Semantic Web data varies in time according to the

size of the data set, however the resource requirements are generally low. Whether this performance can be maintained with significantly larger data sets should be investigated in future research. It may transpire that modifications are required to focus the algorithms on computing a smaller number of more highly relevant relationships. Both this issue and that of scale from a user perspective require future research on a data set perhaps two orders of magnitude greater, in order to be properly informative.

The final issue of scale impacts at a more technical level. Research Question 5 asked *"How feasible is the implementation of user-oriented systems that exploit such algorithms?"* The development and deployment of Hoonoh demonstrates that it is feasible; however a number of technical issues did complicate the process and call aspects of the Semantic Web-based approach into question, on a purely practical level. Section 7.3 describes the changes that were made to the implementation of the algorithms to overcome performance issues. The result was that a relational cache was used to dramatically reduce computation time, whereas the bottleneck remains in generating RDF statements from this relational cache and inserting them into the live Hoonoh triplestore.

No doubt the use of an enterprise class RDF store would mitigate this issue; however the performance of current technologies that can be run in cheap, shared hosting environments is inadequate for handling large datasets, especially when compared to more well-established technologies such as the MySQL database server. This, along with ensuring the quality of freely available libraries for handling RDF, presents a major challenge to the Semantic Web community if adoption is to be seen among traditional developer communities.

9.3.3. Reliance on Proxy Metrics

The empirical study reported in Chapter 4 identified five factors that influence choice of sources and their perceived trustworthiness, of which the most significant three were used in the remainder of the research: experience, expertise and affinity. Data is not readily available from which experience and expertise metrics can be derived (as described in Chapter 7), consequently proxy metrics of usage and credibility were developed for these factors respectively. Whilst these proxy metrics provide a reasonable foundation on which to conduct this research, future work should examine data sources from which direct measures of experience and expertise can be derived.

The work of authors such as Zhu, Song, Rueger et al. (2006) may provide additional means for mining expertise relations between people and topics. However, such relations may be a less robust measure of expertise than the credibility metrics developed in this research, as they are based simply on co-occurrence of terms and names in Web documents. Therefore despite being labelled as an 'expert finding' application this might be better described as 'experience finding', a criticism that applies to much of the work in the expert finding domain.

An alternative approach that should be explored in future research is the harvesting of robust expertise data from accreditation bodies and similar organisations. For example, reputable organisations who certify the knowledge and qualifications of tradespeople and professionals (such as doctors or plumbers) may choose to make these certifications available on the Web in machine-readable form. This would then allow expertise data to be harvested from trusted sources for use in applications such as Hoonoh. How the

trustworthiness of these sources might be determined automatically or on a large scale remains another question for future research.

Whilst the availability of such information may remove the need for proxy metrics for expertise, there is still be a case for including credibility metrics in Hoonoh. Whilst qualifications can serve as a good indicator of the potential for expertise, this is unlikely to arise without endorsement by others. Furthermore, in many cases a position of expertise can be reached without formal qualifications in the area.

Examination of the themes identified by the study in Chapter 4 demonstrates that the expertise factor does subsume themes related to both qualification and credibility (see Appendix B). This may indicate that expertise is best captured by combining metrics that capture both these themes, or that expertise should be replaced by separate qualification and credibility factors. Further research is required to investigate this issue.

9.3.4. Exploiting Semantics for Trust Generation

Tagging with more Semantics

At the time that Revyu was developed a usable yet comprehensive reference source was not available for common concepts or terms with which people may wish to tag reviewed items. Wordnet (van Assem, Gangemi et al., 2006) was considered as such a reference source, however it was deemed too complex to be used by non-specialists as a basis for semantically tagging reviewed items. Use of classes from large ontologies (or classes from many small ontologies) was rejected on the same basis. For these reasons it was decided to simply use unstructured keyword tagging within Revyu and mint a URI for each keyword used in the system to support identification of topics and integration with

other data sources. This approach has the benefit of simplicity, but suffers from limitations caused by synonymy and polysemy.

Future work should investigate the use of URIs from major reference sources such as *DBpedia* and *Geonames* to identify specific concepts, topics or items, in place of URIs minted from keyword tags. For example, the tag 'restaurant' applied to a reviewed item could be identified by the *DBpedia* URI <http://dbpedia.org/resource/Restaurant> rather than <http://revyu.com/tags/restaurant>. This would instantly improve the potential for semantic interoperability of Revyu data with that from other sources that reference *DBpedia* URIs without requiring mappings to be created. Abstracts of *Wikipedia* articles and disambiguation information could be used to support users in choosing appropriate senses where homonyms exist. A similar approach could be adopted using place names and the corresponding URIs from *Geonames*.

The logical extension of this 'semantic tagging' would be to create an interface that did allow users of Revyu to categorise or otherwise annotate reviewed items in a semantically unambiguous fashion. The goal of such a development would be to increase the level of semantics present in Revyu (and ultimately Hoonoh) without substantially increasing the load on the user or requiring additional techniques for deriving semantics from tagging data. For example, the Revyu reviewing form may be extended to ask users to state (in the form of a keyword tag, in separate fields) the 'type' of the item they are reviewing and where appropriate its location. Remaining tags the user wished to apply could be entered into a free text field in the manner of the current Tags field (see Figure 10).

A less user-intensive approach would be to retrieve semantic descriptions of reviewed items from the Web, as these become more widely available, and use these as the basis for computing topical trust metrics in Hoonoh. This would reduce load on the user, remove issues that stem from the syntactic nature of tagging, and broaden the scope of generated trust metrics by exploiting semantic structures already defined on the Web.

Semantic Propagation of Trust Relationships

Grounding Hoonoh topics in formal semantic structures rather than tags creates many possibilities for propagating trust relationships to semantically related concepts. For example, someone who has reviewed many items located in Paris could reasonably be said to have some experience of the topic 'France', irrespective of whether they (or other people reviewing the same items) have used that tag. Using *Geonames* to identify larger regions in which reviewed items are located would enable semantic propagation of trust relationships in this fashion.

Similarly, pre-population of Revyu with film data from *DBpedia* would enable all reviewers of a particular item to be assigned an experience or expertise score for concepts related to the film but which they may not have explicitly mentioned. For example, reviewers may not explicitly tag a film with the name of its director or major stars, however retrieving this information from *DBpedia* and generating new person \rightarrow topic trust metrics accordingly would increase the topical coverage in Hoonoh.

These proposed techniques are somewhat analogous to how previous researchers have attempted to propagate trust relationships through social networks. In this case the propagation is semantic rather than social, exploiting relationships expressed on the

Semantic Web. A key issue for future research is determining the weight or value that should be applied to trust relationships derived in this fashion.

Trust Decay

At present the Hoonoh algorithms can be re-executed as frequently as desired in order to generate up-to-date trust metrics. However, the algorithms do not currently take account of the age of reviews in computing trust metrics, and consequently are not sensitive to potential decay in trust relationships. For example, the trustworthiness of an individual as a source of knowledge on ancient history may decay very slowly, whereas trust in another individual as a source of restaurant recommendations in London may quickly decay if it isn't regularly updated.

Future work should investigate how trust relationships may decay over time, and how rates of decay may vary across different trust factors and, in the case of experience and expertise, across different domains. Such investigations are likely to require fundamental empirical studies with humans in order to develop valid models of how the currency of their knowledge affects the trustworthiness of sources. The findings of such research could then be integrated into the next generation of the Hoonoh algorithms.

9.3.5. Combining Trust Metrics

With the potential to use an increasing number of data sources as evidence of trust relationships, the question arises of how best to combine evidence from these disparate sources when generating unified trust metrics. For example, how might one person's reviews of restaurants in London be combined with data from *Flickr* indicating that she has tagged (or 'geotagged') photos 'london' and data from the Semantic Web stating that she works for a company located in the same city, to produce one valid experience metric

for the topic 'London'? Use of further data sources has the potential to vastly increase coverage in Hoonoh, but this example demonstrates that the greater the number of sources used the more sophisticated the methods must be for combining trust evidence. In addition consideration will need to be given as to how to determine the trustworthiness of arbitrary data sources from which evidence may be gathered.

9.3.6. Evaluation Methods

Lastly, future work should consider how best to evaluate and compare systems such as Hoonoh. The method used in this research appears suitable for evaluating a single system, and could be used to evaluate alternative algorithms operating over the same data, but does not necessarily allow for reliable comparison of different systems using different data sets except by simply comparing correlation coefficients.

Standardised data sets against which systems can be evaluated may be a feasible development. However, given that this would require the use of personal data (such as individual identities, social network connections and trust ratings of others), gaining consent for publishing such a data set may prove impractical. The alternative approach is to use data publicly available on the Web; however this may provide insufficient coverage in terms of reviews or insufficient network density, and is generally not accompanied by trust ratings of other individuals who may feature in the evaluation.

9.4. Outlook

The research presented in this dissertation opens many interesting doors for future research, and signposts several avenues for development of Web technologies that are simultaneously more human-oriented and more human-inspired.

By publishing reviews in such a way that they can be combined with and filtered through social networks, Revyu short-circuits the review spam issue by providing both a means to reduce the impact of spurious reviews and a disincentive to their creation in the first instance. If widely adopted, this characteristic combined with the republishing of reviews in an easily reusable form has the potential to significantly improve the value of reviews and ratings on the Web.

Hoonoh adds to this the essential dimension of trust, as a means to increase relevance, reduce information overload and provide a more personalised experience of information-seeking on the Web. At present, these functionalities can only be provided by those with access to vast data sets and significant computational resources. Organisations such as e-commerce sites that collect user profiles and derive co-preference relationships or taste-overlaps between users have little incentive to make this data available for use by competing services, and as a result it remains locked away in closed worlds.

Whether data sets are open or closed has little bearing, from a technical perspective, on the potential utility of the Hoonoh algorithms. For example, e-commerce or social media sites such as *Amazon*, *YouTube*⁴³ or *last.fm* may choose to adopt these algorithms to generate trust metrics based purely on their own extensive data sets, without making any reference to or use of Revyu data, or making the subsequent metrics available to third parties.

However, by deriving these trust metrics from existing data sources and making them publicly available, Hoonoh allows many competing services to use this knowledge in

⁴³ <http://www.youtube.com/>

providing novel functionality to the user. The effect of this may be to democratise recommendation and personalisation services on the Web.

Furthermore, in richness, the Hoonoh trust model of experience, expertise and affinity surpasses the relationships generally derived by systems developed in this area. This is a direct consequence of the grounding of Hoonoh in first-principles research into the role of social networks in human information- and recommendation-seeking. As the Web (of documents and of data) assumes greater and greater importance in everyday lives, deeply embedding this social element in online applications will be of increasing importance and increasing utility.

For example, reviews written on Revyu, and subsequent trust metrics published via Hoonoh, fulfil a social awareness function which has not yet been systematically examined (this would require a longer term study than is feasible in the context of this research). Anecdotal evidence suggests there are several incidental benefits in writing reviews: others can learn about and stay in touch with one's current activities (e.g. following the progress of foreign trips by reading one's restaurant reviews – "How was your trip to X? I'd forgotten you were going until I saw your reviews"); reviews can highlight shared experiences that were not otherwise apparent (e.g. two friends reviewing the same book without previously realising that the other has also read it).

Fully realising the potential of such functionality requires a shift in emphasis from seeing socially-oriented features as add-ons to existing systems, to seeing social functionality as an essential and central component of a system, fully integrated from the outset. The *Social Computing* (Schuler, 1994) and *Social Navigation* (Dieberger, Dourish et al., 2000) research agendas have previously promoted this perspective, aspects of which can

be seen in social networking sites such as *Facebook*. More recently similar ideas have resurfaced under the umbrella labels of the *Social Web* (e.g. Gruber, 2006) and *Social Semantic Web* (Mikroyannidis, 2007). Applying this approach even more rigorously to Hoonoh, it would be interesting to complement the search- and topic-centric aspects of the system with an exploration of the system's potential as a 'social browser', giving access to the entire Web through the perspective of one's social network.

While Hoonoh is a search-oriented system, the underlying metrics may prove valuable in supporting a more exploratory mode of interaction with the Web. For example, an application or browser plugin could be developed that allowed a user to adopt a certain perspective when browsing, such as that of a domain expert or a specific group of individuals with whom they share a strong affinity. This perspective, and the system's knowledge about the individuals of whom it is comprised, could then be used to prioritise, filter or provide additional annotations of Web resources encountered in a browsing session.

Alternatively, in some situations the user may value the greater potential for serendipitous discovery provided by a less focused or directed browsing experience. In such cases Hoonoh-based approaches may be less desirable. However, while undirected browsing experiences may give the impression of greater serendipity, discovery of interesting resources purely by luck or chance is unlikely in practice, as connections in the Web are not randomly distributed. Therefore the challenge of supporting seemingly serendipitous discovery is to identify resources with a high chance of being interesting to the user, but that they may not otherwise come across through their normal activities and interactions. In a Hoonoh context, bringing to the user's attention resources favoured by individuals just beyond his or her social network (but perhaps with whom the user shares

a minimal level of affinity) may give the impression of serendipity while ensuring a reasonable probability of interest in the resource.

At present it is too early to predict how social networking services such as *Facebook* will evolve in the longer term. However, early indications do suggest that natively social environments in which a number of tasks can be performed (e.g. messaging, sharing, notification) have the potential to rival more established platforms such as email. The key trend is the move from fully open environments such as email, to those in which interaction is focused among a pre-selected social group. If this trend continued it would represent a return to more socially-mediated forms of information access, which would in turn require a more developed understanding of the interaction between social networks and Web technology. The research presented in this dissertation anticipates this trend, contributes to our understanding of the role of social processes in information-seeking, and demonstrates how these may be reflected in technical systems.

However, despite the critical importance of ensuring that Web applications are grounded in and sensitive to social processes, this is just one dimension along which the Web must develop if it is to fully adapt to and support the people who use it. Social networks make up just one aspect of the context in which people live their lives; geography, available resources, past experiences and personal preferences may also add to this context (Heath, Dzbor et al., 2005). Rather than emphasising just a 'Social Web', members of the Web community should pursue a 'Contextual Web', in which applications and services acknowledge, support and adapt to the contextual factors that shape our everyday actions, online or offline.

References

- Ahn, J.-w., Brusilovsky, P. and Farzan, R. (2005) Investigating Users' Needs and Behavior for Social Search. In *Workshop on New Technologies for Personalized Information Access (PIA 2005)*, at the 10th International Conference on User Modeling, Edinburgh, UK.
- Ansari, A., Essegaiier, S. and Kohli, R. (2000) Internet Recommendation Systems. *Journal of Marketing Research*, 37 (3), pp.363-375.
- Aquin, M. d., Baldassarre, C., Gridinoc, L., Angeletou, S., Sabou, M. and Motta, E. (2007) Watson: A Gateway for Next Generation Semantic Web Applications. In *Poster session, 6th International Semantic Web Conference (ISWC 2007)*, Busan, Korea.
- Ashri, R., Ramchurn, S. D., Sabater, J., Luck, M. and Jennings, N. R. (2005) Trust Evaluation through Relationship Analysis. In *Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, The Netherlands, pp.1005 - 1011.
- Auer, S., Bizer, C., Lehmann, J., Kobilarov, G., Cyganiak, R. and Ives, Z. (2007) DBpedia: A Nucleus for a Web of Open Data. In *6th International Semantic Web Conference (ISWC2007)*, Busan, Korea.
- Ayers, D. and Heath, T.: Review Vocabulary, v0.2 <http://purl.org/stuff/rev#> (accessed 10 December 2007).
- Balabanovic, M. and Shoham, Y. (1997) Fab: content-based, collaborative recommendation. *Communications of the ACM*, 40 (3), pp.66-72.
- Beckett, D.: RDF/XML Syntax Specification (Revised) <http://www.w3.org/TR/rdf-syntax-grammar/> (accessed 10th December 2007).
- Belkin, N. J. (2000) Helping People Find What They Don't Know. *Communications of the ACM*, 43 (8), pp.58-61.
- Berners-Lee, T.: Linked Data <http://www.w3.org/DesignIssues/LinkedData.html> (accessed 10 December 2007).
- Berners-Lee, T. and Fischetti, M. (2000) *Weaving the Web: The Past, Present and Future of the World Wide Web by its Inventor*, Texere, London.
- Berners-Lee, T., Hendler, J. and Lassila, O. (2001) The Semantic Web. *Scientific American*, 284 (5), pp.34-43.
- Bizer, C., Cyganiak, R. and Gauss, T. (2007) The RDF Book Mashup: From Web APIs to a Web of Data. In *3rd Workshop on Scripting for the Semantic Web, 4th European Semantic Web Conference (ESWC2007)*, Innsbruck, Austria.

- Bizer, C., Cyganiak, R. and Heath, T.: How to Publish Linked Data on the Web <http://sites.wiwiis.fu-berlin.de/suhl/bizer/pub/LinkedDataTutorial/> (accessed 10 December 2007).
- Bizer, C., Heath, T., Ayers, D. and Raimond, Y. (2007) Interlinking Open Data on the Web. In *Demonstrations Track, 4th European Semantic Web Conference (ESWC2007)*, Innsbruck, Austria.
- Bonhard, P., Harries, C., McCarthy, J. and Sasse, M. A. (2006) Accounting for Taste: Using Profile Similarity to Improve Recommender Systems. In *SIGCHI Conference on Human Factors in Computing Systems (CHI2006)*, Montreal, Canada.
- Bonhard, P. and Sasse, M. A. (2005) "I thought it was terrible and everyone else loved it" - A New Perspective for Effective Recommender System Design. In *19th British HCI Group Annual Conference*, Edinburgh, UK, pp.251-261.
- Borgatti, S. P. and Cross, R. (2003) A Relational View of Information Seeking and Learning in Social Networks. *Management Science*, 49 (4), pp.432-445.
- Bradley, K., Rafter, R. and Smyth, B. (2000) Case-Based User Profiling for Content Personalisation. In *International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (AH2000)*, Trento, Italy, pp.62-72.
- Brickley, D. and Miller, L.: FOAF Vocabulary Specification 0.9 <http://xmlns.com/foaf/0.1/> (accessed 10 December 2007).
- Cantador, I., Castells, P. and Bellogín, A. (2007) Modelling Ontology-based Multilayered Communities of Interest for Hybrid Recommendations. In *1st International Workshop on Adaptation and Personalisation in Social Systems: Groups, Teams, Communities (SociUM 2007)*, at the *11st International Conference on User Modelling (UM 2007)*, Corfu, Greece.
- Celma, O. (2006) Foafing the Music: Bridging the semantic gap in music recommendation. In *5th International Semantic Web Conference (ISWC2006)*, Athens, GA, USA.
- Cross, R. and Borgatti, S. P. (2004) The Ties That Share: Relational Characteristics that Facilitate Information Seeking. In *Social Capital and Information Technology*, MIT Press.
- Cross, R., Parker, A., Prusak, L. and Borgatti, S. P. (2001) Knowing What We Know: Supporting Knowledge Creation and Sharing in Social Networks. *Organizational Dynamics*, 30 (2), pp.100-120.
- Cross, R., Rice, R. E. and Parker, A. (2001) Information Seeking in Social Context: Structural Influences and Receipt of Information Benefits. *IEEE Transactions of Systems, Man, and Cybernetics*, 31 (4), pp.438-448.

Davis, I. and Vitiello, E. J.: RELATIONSHIP: A vocabulary for describing relationships between people <http://vocab.org/relationship/> (accessed 10 December 2007).

Dean, M. and Schreiber, G.: OWL Web Ontology Language Reference <http://www.w3.org/TR/owl-ref/> (accessed 10 December 2007).

Denning, P. J. (2006) Infoglut. *Communications of the ACM*, 49 (7).

Dieberger, A., Dourish, P., Höök, K., Resnick, P. and Wexelbat, A. (2000) Social Navigation: Techniques for Building More Usable Systems. *Interactions*, 7 (6), pp.37-45.

Dieberger, A., Höök, K., Svensson, M. and Lonnqvist, P. (2001) Social Navigation Research Agenda. In *SIGCHI Conference on Human Factors in Computing Systems (CHI2001)*, ACM Press, Seattle, USA, pp.107-108.

Ding, L., Finin, T. and Joshi, A. (2005) Analyzing Social Networks on the Semantic Web. *IEEE Intelligent Systems*, 20 (1), pp.86-90.

Dourish, P. and Chalmers, M. (1994) Running Out of Space: Models of Information Navigation. In *HCI94*, Glasgow, UK.

Duhan, D. F., Johnson, S. D., Wilcox, J. B. and Harrell, G. D. (1997) Influences on Consumer Use of Word-of-Mouth Recommendation Sources. *Journal of the Academy of Marketing Science*, 25 (4), pp.283-295.

Eastlake, D. and Jones, P.: RFC 3174 - US Secure Hash Algorithm 1 (SHA1) <http://isc.faq.org/rfcs/rfc3174.html> (accessed 10 December 2007).

Eisenberg, M. and Schamber, L. (1988) Relevance: The Search for a Definition. In *51st Annual Meeting of the American Society for Information Science*, Atlanta, USA.

Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P. and Berners-Lee, T.: Hypertext Transfer Protocol -- HTTP/1.1 <http://www.w3.org/Protocols/rfc2616/rfc2616.html> (accessed 10 December 2007).

Garton, L., Haythornthwaite, C. and Wellman, B. (1997) Studying Online Social Networks. *Journal of Computer-Mediated Communication*, 3 (1).

Gaved, M., Heath, T. and Eisenstadt, M. (2006) Wikis of Locality: Insights from the Open Guides. In *2006 International Symposium on Wikis (WikiSym2006)*, Odense, Denmark.

Giereth, M. (2005) On Partial Encryption of RDF-Graphs. In *4th International Semantic Web Conference (ISWC2005)*, Galway, Ireland.

Golbeck, J. and Hendler, J. (2004) Reputation Network Analysis for Email Filtering. In *1st Conference on Email and Anti-Spam (CEAS2004)*, Mountain View, CA, USA.

- Golbeck, J. and Hendler, J. (2006) FilmTrust: Movie Recommendations using Trust in Web-based Social Networks. In *IEEE Consumer Communications and Networking Conference (CCNC2006)*, Las Vegas, NV, USA.
- Golbeck, J., Parsia, B. and Hendler, J. (2003) Trust Networks on the Semantic Web. In *7th International Workshop on Cooperative Information Agents (CIA2003)*, Helsinki, Finland.
- Goldberg, D., Nichols, D., Oki, B. M. and Terry, D. (1992) Using Collaborative Filtering to Weave an Information Tapestry. *Communications of the ACM*, 35 (12), pp.61-70.
- Granovetter, M. (1983) The Strength of Weak Ties: A Network Theory Revisited. *Sociological Theory*, 1, pp.201-233.
- Granovetter, M. S. (1973) The Strength of Weak Ties. *The American Journal of Sociology*, 78 (6), pp.1360-1380.
- Gruber, T. (2006) Where the Social Web Meets the Semantic Web (Keynote Talk). In *5th International Semantic Web Conference (ISWC2006)*, Athens, GA, USA.
- Gruber, T.: TagCommons <http://tagcommons.org/> (accessed 1st June 2007).
- Guha, R. (2004) Open Rating Systems. In *1st Workshop on Friend of a Friend*, Galway, Ireland.
- Halaschek-Wiener, C., Golbeck, J., Schain, A., Grove, M., Parsia, B. and Hendler, J. (2005) PhotoStuff – An Image Annotation Tool for the Semantic Web. In *Poster Track, 4th International Semantic Web Conference (ISWC2005)*, Galway, Ireland.
- Hammond, T., Hannay, T., Lund, B. and Scott, J. (2005) Social Bookmarking Tools (I): A General Review. *D-Lib Magazine*, 11 (4).
- Haythornthwaite, C. (2001) The Internet in Everyday Life. *American Behavioral Scientist*, 45 (3), pp.363-382.
- Heath, T., Domingue, J. and Shabajee, P. (2006) User Interaction and Uptake Challenges to Successfully Deploying Semantic Web Technologies. In *3rd International Semantic Web User Interaction Workshop (SWUI2006)*, *5th International Semantic Web Conference (ISWC2006)*, Athens, GA, USA.
- Heath, T., Dzbor, M. and Motta, E. (2005) Supporting User Tasks and Context: Challenges for Semantic Web Research. In *Workshop on End-User Aspects of the Semantic Web (UserSWeb2005)*, *2nd European Semantic Web Conference (ESWC2005)*, Heraklion, Crete.
- Herlocker, J., Konstan, J., Terveen, L. and Riedl, J. (2004) Evaluating Collaborative Filtering Recommender Systems. *ACM Transactions on Information Systems*, 22 (1), pp.5-53.

Herlocker, J. L., Konstan, J. A. and Riedl, J. (2000) Explaining Collaborative Filtering Recommendations. In *ACM Conference on Computer Supported Cooperative Work (CSCW2000)*, Philadelphia, USA, pp.241-250.

Hill, W. and Terveen, L. (1996) Using Frequency-of-mention in Public Conversations for Social Filtering. In *ACM Conference on Computer Supported Cooperative Work (CSCW1996)*, Cambridge, MA, USA.

Howell, D. C. (2002) *Statistical Methods for Psychology*, Duxbury, Pacific Grove, CA, USA.

Jacobs, I. and Walsh, N.: Architecture of the World Wide Web, Volume One <http://www.w3.org/TR/webarch/> (accessed 10 December 2007).

Jeh, G. and Widom, J. (2003) Scaling Personalized Web Search. In *12th International Conference on World Wide Web*, Budapest, Hungary.

Johnson Brown, J. and Reingen, P. H. (1987) Social Ties and Word-of-Mouth Referral Behaviour. *Journal of Consumer Research*, 14 (3), pp.350-362.

Josang, A., Ismail, R. and Boyd, C. (2007) A Survey of Trust and Reputation Systems for Online Service Provision. *Decision Support Systems*, 43 (2), pp.618-644.

Kautz, H., Selman, B. and Shah, M. (1997a) The Hidden Web. *AI Magazine*, 18 (2), pp.27-36.

Kautz, H., Selman, B. and Shah, M. (1997b) Referral Web: Combining Social Networks and Collaborative Filtering. *Communications of the ACM*, 40 (3), pp.63-65.

Klyne, G. and Carroll, J. J.: Resource Description Framework (RDF): Concepts and Abstract Syntax <http://www.w3.org/TR/rdf-concepts/> (accessed 10 December 2007).

Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R. and Riedl, J. (1997) GroupLens: Applying Collaborative Filtering to Usenet News. *Communications of the ACM*, 40 (3), pp.77-87.

Kuhltau, C. C. (1991) Inside the Search Process: Information Seeking from the User's Perspective. *Journal of the American Society for Information Science*, 42 (5), pp.361-371.

Leavesley, J. and Davis, I.: Talis Platform: Harnessing Sophisticated Mass Collaboration on a Global Scale http://www.talis.com/platform/resources/assets/harnessing_sophisticated_mass.pdf (accessed 10 December 2007).

Linden, G., Smith, B. and York, J. (2003) Amazon.com Recommendations: Item-to-Item Collaborative Filtering. *IEEE Internet Computing*, 7 (1), pp.76-80.

Loizou, A. and Dasmahapatra, S. (2006) Recommender Systems for the Semantic Web. In *International Workshop on Recommender Systems, at the 17th European Conference on Artificial Intelligence (ECAI 2006)*, Trento, Italy.

Marchionini, G. (1995) *Information Seeking in Electronic Environments*, Cambridge University Press, Cambridge.

Marsh, S. Formalising Trust as a Computational Concept. Stirling: University of Stirling, 1994.

Massa, P. and Avesani, P. (2004) Trust-aware Collaborative Filtering for Recommender Systems. In *International Conference on Cooperative Information Systems (CoopIS)*, Agia Napa, Cyprus.

Massa, P. and Bhattacharjee, B. (2004) Using Trust in Recommender Systems: an Experimental Analysis. In *International Conference on Trust Managment*, Oxford, England.

McBride, B. (2002) Four Steps Towards the Widespread Adoption of a Semantic Web. In *1st International Semantic Web Conference (ISWC2002)*, Sardinia, Italy.

McDonald, D. W. and Ackerman, M. S. (1998) Just Talk to Me: A Field Study of Expertise Location. In *ACM Conference on Computer Supported Cooperative Work (CSCW1998)*, Seattle, USA.

McDonald, D. W. and Ackerman, M. S. (2000) Expertise Recommender: A Flexible Recommendation System and Architecture. In *ACM Conference on Computer-Supported Cooperative Work (CSCW2000)*, Philadelphia, USA.

McPherson, M., Smith-Lovin, L. and Cook, J. M. (2001) Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27, pp.415–44.

Mika, P. (2004) Social Networks and the Semantic Web. In *IEEE/WIC/ACM International Conference on Web Intelligence (WI'04)*, Beijing, China.

Mikroyannidis, A. (2007) Toward a Social Semantic Web. *IEEE Computer*, 40 (11), pp.113-115.

Miles, A. and Brickley, D.: SKOS Core Vocabulary Specification
<http://www.w3.org/TR/swbp-skos-core-spec/> (accessed 1st June 2007).

Miller, B. N., Albert, I., Lam, S. K., Konstan, J. and Riedl, J. (2003) MovieLens Unplugged: Experiences with an Occasionally Connected Recommender System. In *International Conference on Intelligent User Interfaces (IUI2003)*, Miami, FL, USA, pp.263-266.

Mobasher, B., Cooley, R. and Srivastava, J. (2000) Automatic Personalization Based on Web Usage Mining. *Communications of the ACM*, 43 (8), pp.142-151.

Möller, K., Heath, T., Handschuh, S. and Domingue, J. (2007) Recipes for Semantic Web Dog Food - The ESWC and ISWC Metadata Projects. In *6th International Semantic Web Conference (ISWC2007)*, Busan, Korea.

Morrison, E. W. and Vancouver, J. B. (2000) Within-Person Analysis of Information Seeking: The Effects of Perceived Costs and Benefits. *Journal of Management*, 26 (1), pp.119-137.

Narayanan, S., Koppaka, L., Edala, N., Loritz, D. and Daley, R. (2004) Adaptive Interface for Personalizing Information Seeking. *CyberPsychology and Behavior*, 7 (6), pp.683-688.

Newman, R., Russell, S. and Ayers, D.: Tag Ontology
<http://www.holygoat.co.uk/owl/redwood/0.1/tags/> (accessed 10 December 2007).

O'Hara, K., Alani, H., Kalfoglou, Y. and Shadbolt, N. (2004) Trust Strategies for the Semantic Web. In *Workshop on Trust, Security and Reputation on the Semantic Web, 3rd International Semantic Web Conference (ISWC2004)*, Hiroshima, Japan.

Oldakowski, R., Bizer, C. and Westphal, D. (2005) RAP: RDF API for PHP. In *1st Workshop on Scripting for the Semantic Web (SFSW2005), 2nd European Semantic Web Conference (ESWC2005)*, Heraklion, Crete.

O'Reilly, C. A. (1982) Variations in Decision Makers' Use of Information Sources: The Impact of Quality and Accessibility of Information. *The Academy of Management Journal*, 25 (4), pp.756-771.

O'Reilly, T.: What Is Web 2.0
<http://www.oreillyn.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
(accessed 10 December 2007).

Page, L., Brin, S., Motwani, R. and Winograd, T. (1999) The PageRank Citation Ranking: Bringing Order to the Web. Stanford InfoLab.
<http://dbpubs.stanford.edu/pub/1999-66>

Park, T. K. (1994) Toward a Theory of User-Based Relevance: A Call for a New Paradigm of Inquiry. *Journal of the American Society for Information Science*, 45 (3), pp.135-141.

Prud'hommeaux, E. and Seaborne, A.: SPARQL Query Language for RDF
<http://www.w3.org/TR/rdf-sparql-query/> (accessed 10th December 2007).

Ramsey, P. H. (1989) Critical Values for Spearman's Rank Order Correlation. *Journal of Educational and Behavioural Statistics*, 14 (3), pp.245-253.

Resnick, P. and Varian, H. R. (1997) Recommender Systems. *Communications of the ACM*, 40 (3), pp.56-58.

- Richardson, M., Agrawal, R. and Domingos, P. (2003) Trust Management for the Semantic Web. In *2nd International Semantic Web Conference (ISWC2003)*, Florida, USA.
- Salter, J. and Antonopoulos, N. (2006) CinemaScreen Recommender Agent: Combining Collaborative and Content-Based Filtering. *IEEE Intelligent Systems*, 21 (1), pp.35-41.
- Schafer, J. B., Konstan, J. A. and Riedl, J. (2001) E-Commerce Recommendation Applications. *Data Mining and Knowledge Discovery*, 5 (1-2), pp.115-153.
- Schuler, D. (1994) Social Computing. *Communications of the ACM*, 37 (1), pp.29.
- Scott, J. (2000) *Social Network Analysis: A Handbook*, Sage, London.
- Shadbolt, N. R., Hall, W. and Berners-Lee, T. (2006) The semantic Web revisited. *IEEE Intelligent Systems*, 21 (3), pp.96-101.
- Shardanand, U. and Maes, P. (1995) Social Information Filtering: Algorithms for Automating "Word of Mouth". In *SIGCHI Conference on Human Factors in Computing Systems (CHI1995)*, Denver, CO, USA.
- Sinha, R. and Swearingen, K. (2001) Comparing Recommendations Made by Online Systems and Friends. In *DELOS Workshop: Personalisation and Recommender Systems in Digital Libraries*, Dublin, Ireland.
- Sinha, R. and Swearingen, K. (2002) The Role of Transparency in Recommender Systems. In *SIGCHI Conference on Human Factors in Computing Systems (CHI2002)*, Minneapolis, USA, pp.830-831.
- Smith, J. A. (1995) Semi-Structured Interviewing and Qualitative Analysis. In *Rethinking Methods in Psychology*, Sage, London, pp.9-26.
- Specia, L. and Motta, E. (2007) Integrating Folksonomies with the Semantic Web. In *4th European Semantic Web Conference (ESWC2007)*, Innsbruck, Austria.
- Svensson, M., Hook, K., Laaksolahti, J. and Waern, A. (2001) Social Navigation of Food Recipes. In *SIGCHI Conference on Human Factors in Computing Systems (CHI2001)*, Seattle, USA.
- Teevan, J., Dumais, S. T. and Horvitz, E. (2005) Beyond the Commons: Investigating the Value of Personalized Web Search. In *Workshop on New Technologies for Personalized Information Access (PIA2005)*, Edinburgh, Scotland.
- Terveen, L., Hill, W., Amento, B., McDonald, D. and Creter, J. (1997) PHOAKS: A System for Sharing Recommendations. *Communications of the ACM*, 40 (3), pp.59-62.
- Tummarello, G., Oren, E. and Delbru, R. (2007) Sindice.com: Weaving the Open Linked Data. In *6th International Semantic Web Conference (ISWC2007)*, Busan, Korea.

van Assem, M., Gangemi, A. and Schreiber, G.: RDF/OWL Representation of WordNet <http://www.w3.org/2001/sw/BestPractices/WNET/wn-conversion.html> (accessed 10 December 2007).

van Setten, M., McNee, S. M. and Konstan, J. A. (2005) Beyond Personalization: The Next Stage of Recommender Systems Research. In *International Conference on Intelligent User Interfaces (IUI2005)*, San Diego, USA.

Völkel, M., Krötzsch, M., Vrandečić, D., Haller, H. and Studer, R. (2006) Semantic Wikipedia. In *15th International Worldwide Web Conference*, Edinburgh, UK.

W3C Technical Architecture Group (TAG): [httpRange-14: What is the range of the HTTP dereference function?](http://www.w3.org/2001/tag/issues.html#httpRange-14) <http://www.w3.org/2001/tag/issues.html#httpRange-14> (accessed 10 December 2007).

Walsh, G. and Swinford, S. (2006) Hotel review websites: a five-star scam. In *The Sunday Times*, London. 12 November 2006.

Wexelblat, A. and Maes, P. (1999) Footprints: History-Rich Tools for Information Foraging. In *SIGCHI Conference on Human Factors in Computing Systems (CHI1999)*, Pittsburgh, USA, pp.270-277.

Whyte, W. H. W. J. (1954) The Web of Word of Mouth. *Fortune*, November.

Yaniv, I. and Kleinberger, E. (2000) Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation. *Organizational Behavior and Human Decision Processes*, 83 (2), pp.260-281.

Zhu, J., Song, D., Rueger, S., Eisenstadt, M. and Motta, E. (2006) The Open University at TREC 2006 Enterprise Search Track Expert Search Task. In *Fifteenth Text REtrieval Conference (TREC2006)*, Gaithersburg, MD, USA.

Ziegler, C.-N. and Lausen, G. (2004) Analyzing Correlation between Trust and User Similarity in Online Communities. In *2nd International Conference on Trust Management (iTrust2004)*, Oxford, England.

Appendices

Appendix A: Interview Script

Preamble (Experimenter's Prompts)

1. Disclosure: Data will all be anonymised. The study can be stopped at any time for any reason and all data destroyed, you just have to ask.
2. I'm carrying out a study into how people use recommendations from those around them to help carry out tasks or solve problems. I will read you some scenarios and ask you some questions about how you would approach the problems in these scenarios.
3. Please answer the questions in as much detail as possible, explaining the reasons for any decisions you make. There are no wrong answers; I'm just interested in how you approach the problem.

Scenarios

Scenario 1

You move into a new house that requires renovation, including some substantial plumbing work.

- Who would you ask about recommended plumbers? Please be specific and cite individuals or groups of people if possible. Why would you ask this person/group?
- Is there anyone you wouldn't ask? Why not?
- If the first person or group you sought the recommendation from was not available, who would you ask next? Why?
- If this recommendation turned out to be a poor one, what effect would that have on your seeking of recommendations from that person or group in the future in the same topic area? What about in different areas?

Scenario 2

You are travelling to Madrid on business and need to find a hotel to stay in during your visit.

- Who would you ask about recommended hotels? Please be specific and cite individuals or groups of people if possible. Why would you ask this person/group?
- Is there anyone you wouldn't ask? Why not?
- If the first person or group you sought the recommendation from was not available, who would you ask next? Why?
- If this recommendation turned out to be a poor one, what effect would that have on your seeking of recommendations from that person or group in the future in the same topic area? What about in different areas?

Scenario 3

You are suffering from moderate and ongoing back pain and need to find some ways of getting it treated.

- Who would you ask about recommended ways of getting it treated? Please be specific and cite individuals or groups of people if possible. Why would you ask this person/group?
- Is there anyone you wouldn't ask? Why not?
- If the first person or group you sought the recommendation from was not available, who would you ask next? Why?
- If this recommendation turned out to be a poor one, what effect would that have on your seeking of recommendations from that person or group in the future in the same topic area? What about in different areas?

Scenario 4

You are planning a holiday to the east coast of the USA and need to find some information about how to spend your time there.

- Who would you ask about recommended activities? Please be specific and cite individuals or groups of people if possible. Why would you ask this person/group?
- Is there anyone you wouldn't ask? Why not?
- If the first person or group you sought the recommendation from was not available, who would you ask next? Why?
- If this recommendation turned out to be a poor one, what effect would that have on your seeking of recommendations from that person or group in the future in the same topic area? What about in different areas?

Prompt questions if participant has real examples of their own

- 1) What was the information you needed?
- 2) Who did you ask, and why specifically did you choose them?
- 3) Were they able to give an answer?
- 4) If so, did it prove useful? If not, what did you do next?
- 5) How did that outcome influence your willingness to seek/accept advice from them in the future?

Appendix B: Master list of themes and superordinate themes identified in the interview study

1. People you would ask Why would you ask them?
Themes : expertise, local knowledge, mindset, experience, quality, history, shared interests, similarity, knowledge, quantity, language used, closeness (relationship-wise), respect, track record, trust, blameable, contacts, proximity (physical), availability, faith, ability, range (diversity), validated, insight, taken seriously, helpfulness, gatekeeper, diversity (of opinion), qualification, appropriateness, knowledge of me, appearance, relevance, authority, suitability, similar needs, comparability, similar taste, existing bandwidth, standards (similar), similar values, judgement, viewpoint, individuality, reliability (accuracy), shared background, outlook, length of knowing, ease of use, specialism, authority, endorsement/validated locally, respect opinions, standards, preferences, shared likes, like me, like the sound of them, personal taste.
2. People you would not ask Why would you not ask them?
Themes: vested interest, inappropriate (socially), no knowledge, no experience, poor track record, biased, non-constructive, language style, different expectations, poor gatekeeper, gatekeeper to a poor solution, unqualified, inappropriate source for me, different lifestyle, inappropriate to your needs, difference in wealth, different priorities, different interests, infrequent contact, not comfortable with, unreliable, untrustworthy, untailored to me, no info about source, no recommendation, different background, different values, company that wasn't liked, non-specialist

3. Grouped Themes
(the three headings of trust, practicalities and diversity emerged from the data)
Trust <ul style="list-style-type: none">- Knows about these things/May know something about- Is seen as an expert in, is an authority in- Has had experience of- Is like me in some way: standards, viewpoint, outlook, values, background, interests, approach, mindset, taste, judgement, preferences, shared likes, expectations, lifestyle, needs, priorities, wealth- Has given good recommendations in the past- Is qualified in- Knows me/would give appropriate solution- I am close to them/I am comfortable with them- Is seen as expert in- Has vested interest in/Is biased
Practicalities <ul style="list-style-type: none">- Is easy to ask/available/- Has lots of contacts- Would be helpful/constructive- Is not appropriate to ask
Diversity <ul style="list-style-type: none">- There are lots of them, wide range- Would give interesting solutions to

4. Superordinate trust themes

(based on a further grouping of the first set of groupings)

Track Record

- has given good recommendation in the past

Affinity

- is like me in some way
- appropriate solution

Impartiality

- is impartial

Experience Of

- has experience of

Has Expertise In (either acquired or validated)

- knows about these things
- is seen as an expert in
- is qualified in